A Relational Complexity Approach to the Development of Hot/Cool Executive Functions

Katie Bunch
Bachelor of Psychology (Honours)

School of Psychology,
Griffith University, Gold Coast

Submitted in fulfillment of the requirements of the degree of a

Doctor of Philosophy

August, 2006
ABSTRACT

Previous research indicates that many important changes in executive functions, or higher cognitive capacities, occur between the ages of three and five years. Additionally, a distinction can be made between the cognitive functions associated with two different cortical regions. The functions of the dorsolateral prefrontal cortex (DL-PFC) are assessed using ‘cool’ tasks that are abstract and decontextualised. In contrast, the functions of the orbitofrontal cortex (OFC) are assessed using ‘hot’ tasks that require flexible appraisal of the affective significance of stimuli (Zelazo & Müller, 2002). Different clinical populations have been hypothesized to differ in terms of their impairment on tasks associated with each area of functioning. Current research conclusions regarding the primacy of hot versus cool executive function impairments are limited, however, as they have not taken complexity into account. That is, tasks currently used in investigations of hot and cool executive functions might differ in terms of the complexity of the cognitive processes that the tasks require. Therefore, comparisons across tasks may be misleading because these tasks vary in terms of the demands they place on participants as well as their hot versus cool status.

While complexity theories have been applied to a number of cool tasks, only one hot task, those measuring theory-of-mind abilities, have been analysed in terms of complexity. One aim of the current research was to modify several tasks presumed to measure OFC performance to include a complexity manipulation. Tasks from three hot domains (conditional discrimination, the Children’s Gambling Task, and future-oriented decision-making) were analysed in terms of their relational complexity, that is, the number of related entities or arguments inherent in a task or concept (Halford, 1993). Based on these complexity analyses, binary-relational and ternary-relational items of each of these tasks were developed or existing tasks were selected and/or modified. The binary-relational
items were closely matched to the ternary-relational items in terms of stimuli and procedure, however, they were lower in complexity.

After pilot testing, the three new measures of hot executive functioning were included in a larger test battery that was administered to a sample of 120 normally developing 3-, 4-, 5-, and 6-year-old children. Existing binary- and ternary-relational items assessing theory-of-mind (a hot task) and three cool measures (transitivity, class inclusion and the Dimensional Change Card Sort test) were also included. The inclusion of measures of both hot and cool executive functions, each with complexity manipulated, allowed for the examination of a possible differential age of emergence of executive abilities associated with the DL-PFC versus the OFC.

In support of the relational complexity approach, significant complexity effects were found across all seven tasks. Items at a higher level of complexity were experienced as relatively more difficult by children of all ages. Significant effects of age were also observed, with performance across all tasks increasing with age. The age effects were strongest on the ternary-relational items. The pass-fail data indicated that the majority of children in all age groups succeeded on the binary-relational items. However, it was not until a median of five years of age that children were able to process ternary relations. Consequently, the ternary-relational items produce the greatest differences in performance between the four age groups.

The overall pattern of the results also suggested that a distinction can be made between the ages of emergence of abilities associated with the OFC versus the DL-PFC. The results of the pass-fail percentages, patterns of age-related change and age effects on domain factor scores all suggested that while hot executive functions may begin to develop around four years of age, similar levels of improvement are not seen in cool executive functions until five years of age. Thus, the ability to succeed on ternary-relational items of
hot executive function tasks appeared to emerge slightly earlier than the cool executive function tasks.

Complexity appears to be a critical factor underlying children’s performance on executive function tasks, and future assessment regarding the development of executive abilities will benefit from keeping this in mind. While some refinement of new task items may be beneficial, the current test battery may have utility in further examinations of the executive profiles underlying clinical groups, such as children with autism and ADHD.
STATEMENT OF ORIGINALITY

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

___________________________
Katie Bunch
# TABLE OF CONTENTS

## CHAPTER 1. INTRODUCTION TO EXECUTIVE FUNCTIONS

- Executive Functions and the Frontal Lobes ............................................ 2
- Current Approaches to Executive Functions in Adults ............................ 5
  - *Neuropsychological Approach to the Assessment of Executive Functions* ........ 6
  - *Evaluation of the Neuropsychological Approach* .................................. 11
  - *Working Memory Approach* ................................................................. 13
  - *Evaluation of the working memory approach* ........................................ 17
  - *Inhibitory Approach* ............................................................................... 17
  - *Evaluation of the inhibitory approach* .................................................. 20
- *Multiple Component Models* ................................................................. 21
  - Conclusions .............................................................................................. 23

## CHAPTER 2. DEVELOPMENT OF EXECUTIVE FUNCTIONS

- Biological Development of the Frontal Lobes .......................................... 25
- Executive Function and Cognitive Development ........................................ 27
  - *Piagetian Stage Theory* ........................................................................... 27
  - *Neo-Piagetian Theories* .......................................................................... 28
  - *Complexity Approaches to Cognitive Development* .............................. 30
- Delineating Functions of the Prefrontal Cortex ......................................... 35
- Approaches to the Development of Executive Functions in Childhood ........ 40
  - *Global Assessment of Executive Functions in Children* ....................... 41
  - *Refinement of Global Tasks* ................................................................. 43
  - *Working Memory Approach to Development of Executive Function in Children* .......................... 44
  - *Inhibitory Approach to Development of Executive Function in Children* .......................... 47
  - *Multiple Component Approaches* .......................................................... 49
  - *Relational Complexity Approach to Executive Function* ...................... 53
- Conclusions .............................................................................................. 54

## CHAPTER 3. OVERVIEW OF THE EMPIRICAL RESEARCH

- General Method .......................................................................................... 56
- Participants .................................................................................................. 58
- Representativeness of the Sample .............................................................. 59
- Procedure ................................................................................................... 61
- Treatment of Results .................................................................................. 62
CHAPTER 4. COMPLEX DECISION-MAKING ...................................................... 64
Gambling Tasks in Child Populations .............................................................. 66
Complexity Analysis of the Children’s Gambling Task .................................. 67
Method ........................................................................................................ 70
  Materials ................................................................................................. 70
  Procedure ............................................................................................... 72
Results ......................................................................................................... 73
  Pilot Study .............................................................................................. 73
  Main Study ............................................................................................. 78
Discussion ................................................................................................. 83

CHAPTER 5. FUTURE-ORIENTED DECISION-MAKING ................................. 86
Relation Of Future-Oriented Decision-Making to the OFC .............................. 90
Complexity Analysis of Future-Oriented Decision-Making Tasks .................. 91
Method ....................................................................................................... 93
  Materials ............................................................................................... 93
  Procedure ............................................................................................. 93
Results ....................................................................................................... 95
  Pilot Study ............................................................................................ 95
  Main Study .......................................................................................... 100
Discussion ............................................................................................... 104

CHAPTER 6. CONDITIONAL DISCRIMINATION / REVERSAL LEARNING 108
Relation of Object Reversal to the OFC ......................................................... 108
Development of Object Reversal and Conditional Discrimination Abilities ...... 109
Complexity Analysis of Conditional Discrimination / Reversal Learning ........ 111
Method ..................................................................................................... 112
  Conditional Discrimination Materials and Procedure: ............................... 112
  Reversal Learning Stimuli and Procedure ................................................ 115
Results ..................................................................................................... 117
  Pilot Study ............................................................................................ 117
  Main Study .......................................................................................... 122
Discussion ............................................................................................... 128
CHAPTER 10. CONCLUSIONS ........................................................................... 186

Theoretical Implications .................................................................................. 186

Complexity Theories ......................................................................................... 186

Hot-Cool Distinction ......................................................................................... 188

Methodological Issues ....................................................................................... 189

Future Research ............................................................................................... 191

REFERENCES .................................................................................................... 193

APPENDICES ..................................................................................................... 215

Appendix A: Information Letters and Informed Consent ................................. 215

Appendix B: Response Sheet for the Gambling Task ........................................... 217

Appendix C: Response Sheet for Future-Oriented Decision-Making Task ........... 218

Appendix D: Response Sheet for the Transformation and Connections Tasks ...... 219

Appendix E: Response Sheet for Appearance-Reality and False-Belief Tasks ..... 220

Appendix F: Response Sheet for Transitivity ....................................................... 221

Appendix G: Response Sheet for Class Inclusion ................................................. 222

Appendix H: Standards for Templates and Objects Sorted in the Binary and Ternary-
Relational DCCS ................................................................................................. 223

Appendix I: Response Sheets for the DCCS ....................................................... 2234
LIST OF TABLES

Table 3.1 Selection of Tasks in Pilot and Main Study from Hot and Cool Domains of Executive Function ............................................................. 57

Table 3.2 Mean (SD) Performance on NEPSY Subtests Across Age Groups (N = 120) .... 60

Table 3.3 Mean (SD) Executive Domain Scores and Percentile Rankings Across Age Groups in the Main Study .......................................................... 61

Table 4.1 Frequency of Children in Each Age-Group Performing above Chance Levels on the Combined Binary-Relational and Ternary-Relational Versions in the Pilot Study (N = 72) .......................................................................................... 77

Table 4.2 Crosstabulation of Children Passing and Failing the Combined Binary- and Ternary-Relational Versions of the Gambling Task in the Pilot Study (N = 72). 78

Table 4.3 Frequency of Children in Each Age-Group Performing above Chance Levels on the Combined Binary- and Ternary-Relational Versions in the Main Study (N = 120) ............................................................................................. 82

Table 4.4 Crosstabulation of Children Passing and Failing the Combined Binary- and Ternary-Relational Versions in the Main Study (N = 120) ............................................................................................. 83

Table 5.1 Example Binary- and Ternary-Relational Items in the Future-Oriented Decision Making Task ........................................................................ 92

Table 5.2 Binary- and Ternary-Relational Items in the Future-Oriented Decision-Making Task (correct responses are highlighted) .................................................. 95

Table 5.3 Frequency of Children Performing above Chance Levels on the Two Levels of Complexity of the Future-Oriented Decision-Making Task, Pilot Study (N = 72) .................................................................................. 99

Table 5.4 Crosstabulation of Children Passing and Failing the Binary- and Ternary- Relational Items of the Future-Oriented Decision-Making Task in the Pilot Study (N = 72) .................................................................................. 99

Table 5.5 Frequency of Children Performing above Chance Levels on the Two Levels of Complexity of the Future-Oriented Decision-Making Task in the Main Study (N = 120) ........................................................... 103

Table 5.6 Crosstabulation of Children Passing and Failing the Binary- and Ternary- Relational Items of the Future-Oriented Decision-Making Task (N = 120) ..... 103

Table 6.1 Conditional Discrimination Stimuli and Reward Contingencies ............... 113

Table 6.2 Number of Trial Blocks Administered in the Pilot and Main Study of the Conditional Discrimination Task ........................................................................ 115
Table 6.3 Reversal Learning Stimuli and Reward Contingencies........................... 115

Table 6.4 Number of Trial Blocks Administered in the Learning and Reversal Phases of
Problem 1 and 2 of the Reversal Learning Task ...................................................... 117

Table 6.5 Frequency of Children Within Each Age-Group Performing above Chance Levels
on the Binary- and Ternary Relational Items of the Learning Phase (N = 72)..... 121

Table 6.6 Crosstabulation of Children Passing and Failing the Binary- and Ternary-
Relational Items of the Conditional Discrimination Task Learning Phase (N =
72)............................................................................................................................ 122

Table 6.7 Frequency of Children Within Each Age-Group Performing above Chance Levels
on the Reversal Learning and Conditional Discrimination Items (N = 120)..... 127

Table 6.8 Crosstabulation of Children Passing and Failing the Reversal Learning and the
Conditional Discrimination Items (N = 120).............................................................. 127

Table 7.1 Frequency of Children within Each Age Group Performing above Chance Levels
on the Binary-Relational and Ternary-Relational Theory-of-Mind items (N = 120)
............................................................................................................................ 148

Table 7.2 Crosstabulation of Children Passing and Failing the Binary- and Ternary-
Relational Theory-of-Mind Items (N = 120).............................................................. 148

Table 8.1 Frequency of Children Within Each Age-Group Performing above Chance Levels
on the Binary- and Ternary-Relational Items of Transitivity (N = 120)............. 157

Table 8.2 Crosstabulation of Children Passing and Failing the Binary- and Ternary-
Relational Items of Transitivity (N = 120).............................................................. 158

Table 8.3 Composition of Displays used in Class Inclusion Task ......................... 161

Table 8.4 Frequency of Children Within Each Age-Group Performing above Chance Levels
on the Binary- and Ternary-Relational Items of Class Inclusion (N =120) ....... 164

Table 8.5 Crosstabulation of Children Passing and Failing the Binary-Relational and
Ternary-Relational Items of Class Inclusion (N = 120) ........................................ 165

Table 8.6 Frequency of Children in Each Age-Group Performing above Chance Levels on
the Binary- and Ternary-Relational Items of the DCCS (N = 120) ...................... 172

Table 8.7 Crosstabulation of Children Passing and Failing the Binary- and Ternary-
Relational Items of the DCCS (N = 120)................................................................. 173

Table 9.1 Patterns of Age-Related Changed revealed by Scheffe tests on the Ternary-
Relational Items (N = 120)..................................................................................... 176

Table 9.2 Percentage of Children in Each Age Group Passing Ternary-Relational Items. 177
Table 9.3 Zero-order (below diagonal) and Age-partialled (above diagonal) Correlations Among the Three Hot Executive Measures in the Pilot Study \( (N = 72) \) .......... 180

Table 9.4 Mean Composite Score, Zero-order and Age-partialled Correlations Among the Executive Measures and Age \( (N = 120) \) ........................................................................ 181

Table 9.5 Loadings of the Hot Executive Function Factor ......................................................... 183

Table 9.6 Loadings of the Cool Executive Function Factor ......................................................... 184
LIST OF FIGURES

Figure 1.1. Major sub-divisions of the prefrontal cortex (Source: Dempster & Corkill, 1999) ........................................................................................................................................ 2

Figure 1.2. Cytoarchitectonic map of the (a) lateral surface and (b) medial surface of the human cerebral cortex by Brodmann (Source: Kain & Perner, 2005). ...................... 3

Figure 1.3. Material used in the Wisconsin Card Sorting Task . Correct sorting of the response card by the colour criterion would be location (a), correct sorting by number would be location (b), or correct sorting by shape would be location (c). 7

Figure 1.4. Apparatus used in the Tower of Hanoi Task .......................................................... 9

Figure 2.1. Hierarchical tree structure depicting formal relation among rules (Zelazo & Frye, 1998). Note: s₁ and s₂ = setting conditions; a₁ and a₂ = antecedents; c₁ and c₂ = consequences................................................................. 31

Figure 4.1. Sample card face from the ternary-relational disadvantageous deck of the Children’s Gambling Task (gain = 2, loss = -4). ..................................................... 70

Figure 4.2. Mean number (standard error) of advantageous choices on the binary-relational decks (a) and ternary-relational deck (b) of the pilot study (N = 72). ................. 74

Figure 4.3. Mean number (standard error) of advantageous choices on the binary-relational versions (a) and the ternary-relational versions (b) of the pilot study (N = 72). .. 76

Figure 4.4. Mean number (standard error) of advantageous choices on the combined binary-relational versions (a) and the ternary-relational version (b) across the blocks of 10 trials in the main study (N = 120). ................................................................. 80

Figure 5.1. Mean number of correct choices for each age group on the binary- and ternary-relational items of the future-oriented decision-making task, pilot study (N = 72). ........................................................................................................ 96

Figure 5.2. Mean number of correct responses across the three item types and two levels of complexity in the pilot study (N = 72). ................................................................. 97

Figure 5.3. Mean number of correct choices for males and females for the three item types in the pilot study (N = 72). .................................................................................... 98

Figure 5.4. Mean number of correct responses for the four age groups for the binary- and ternary-relational items of the future-oriented decision-making task, main study (N = 120). ........................................................................................................ 101

Figure 5.5. Mean number of correct responses for the three item types and the two levels of complexity in the main study (N = 120). ................................................................. 102

Figure 6.1. Example stimulus slide in the conditional discrimination task ......................... 112
Figure 6.2. Cumulative number of children who met criteria on each trial block on conditional discrimination task (percentage of children who met criteria by block 6 of problem 1; N = 72). .......................................................... 118

Figure 6.3. Cumulative number of children who reached the learning criteria on each block of the learning and reversal phases of problem 1 and 2 (N = 72)...................... 119

Figure 6.4. Mean number of correct choices on the reversal learning (trial block 1 of learning trials and trial block 1 of reversal trials) and conditional discrimination items (trial blocks 1 and 2), problems 1 and 2 combined (N = 72). ..................... 120

Figure 6.5. Cumulative number of children who met criteria on each trial block on conditional discrimination task (percentage of children who met criteria by block 6). ............................................................................................................. 123

Figure 6.6. Cumulative number of children in each age group who reached the learning criteria on the learning and reversal problems (N = 120).......................... 124

Figure 6.7. Mean number correct on the reversal learning and conditional discrimination items of the conditional discrimination as a function of age (N = 120). .......... 125

Figure 7.1. Relational complexity analysis of appearance-reality, false-belief, connections and transformation tasks. .......................................................................................................................... 139

Figure 7.2. Mean (standard error) percentage of items correct for each age-group on the connections/transformation items and appearance-reality/false-belief items (N = 120). ............................................................................................................. 146

Figure 8.1. Example premise display for binary and ternary-relation items for transitivity. .......................................................................................................................... 154

Figure 8.2. Mean performance of the four age groups for the binary- and ternary-relational items of transitivity (N = 120). .......................................................................................................................... 156

Figure 8.3. Mean performance of the four age groups across the binary- and ternary-relational items of the class inclusion task (N = 120). .............................. 163

Figure 8.4. Stimuli for the Dimensional Change Card Sort task ............................................. 168

Figure 8.5. Example template and objects sorted in DCCS task ........................................... 169

Figure 8.6. Mean performance of the four age groups for the binary- and ternary-relational items of the DCCS (N = 120). ................................................................. 171

Figure 9.1. Age effects on the hot and cool domain factor scores. ........................................... 185
ACKNOWLEDGEMENTS

I would like to acknowledge and thank the following:

- Dr Glenda Andrews for her advice, guidance and support through all stages of the proposal, research and draft copies.
- Professor Greame Halford for his input and suggestions on theoretical matters and final drafts.
- The Applied Cognitive and Neuroscience Research Center for their assistance and presentation feedback.
- The staff and students of Griffith University, Gold Coast for their support and encouragement.
- The children, parents and teachers of the participating day-care centers and schools for their cooperation and patience.
- Finally, much love and gratitude to my friends and family (including furry creatures) for their continued encouragement and perseverance through every stage of the PhD process. Special mention to the following:
  - Mum – for her support in every sense of the word.
  - Dad – for believing that I was capable of anything, and everything.
  - Mark – for providing such amazing footsteps to follow.
  - Jo – for being there for the laughter, the tears, and everything in between.
CHAPTER 1. INTRODUCTION TO EXECUTIVE FUNCTIONS

The term executive function refers to a complex cognitive construct, whose theoretical definition remains provisional despite a recent flourish in research. It refers to higher-order cognitive capacities that are at the most superordinate level in the cognitive hierarchy (Tranel, Anderson, & Benton, 1994). Executive functions encompass the skills necessary for goal directed behaviour, the control of cognition, and the regulation of behaviour and thought (Phillips, 1997). They are described as domain general processes that are conscious and effortful. Executive processes can therefore be contrasted with non-executive processes, which are often automatic and modular, and do not involve conscious effort.

Intact executive functions enable human beings to develop and execute plans, form analogies, obey social rules, adapt to unexpected circumstances and perform tasks simultaneously (Grafman & Litvan, 1999). They are crucial to everyday functioning and are important determinants of success at school and in the workforce (Lyon, 1996). In contrast, executive dysfunction is characterized by a marked difficulty in novel or ambiguous situations (Hughes, Leboyer, & Bouvard, 1997), poor planning and reasoning (Anderson, 2002), and difficulty controlling impulsive behaviour (Brophy, Taylor, & Hughes, 2002). It also produces difficulties in understanding the consequences of one’s behaviour on others, in holding abstract ideas and future goals in mind, in inhibiting inappropriate behaviour, and in responding adaptively to changes in the social environment (Brophy et al., 2002). Executive dysfunction has been implicated in several developmental disorders including autism, attention deficit hyperactivity disorder (ADHD) and conduct disorders (e.g., Hughes, 2002; Ozonoff & Jensen, 1999; Pennington & Ozonoff, 1996).
Traditionally, executive functions have been thought to be subserved by the prefrontal cortical region of the brain (Dowsett & Livesey, 2000). The frontal lobes comprise 24 to 29% of the entire surface of the human cerebral cortex and are involved in many aspects of cognitive activity that are considered to be distinctly human (Goldman-Rakic, 1987). Each frontal lobe can be divided into three major subareas: (a) the primary motor cortex, which provides the main cortical output for involuntary movements; (b) the premotor area, involved in the integration and programming of sequential movements; and (c) the prefrontal cortex (Dempster & Corkill, 1999). The prefrontal area is further divided into five major subdivisions (Cummings, 2001). As can be seen in Figure 1.1, these are supplementary motor area, the frontal eye fields, the orbito-frontal cortex (OFC), the dorsolateral prefrontal cortex (DL-PFC), and anterior cingulate-medial frontal cortex (not illustrated). The latter three regions are thought to be particularly relevant to executive functions (Cummings, 2001).

*Figure 1.1. Major sub-divisions of the prefrontal cortex (Source: Dempster & Corkill, 1999).*
The OFC can be dissociated from the DL-PFC by examining the cortical architectonics (Stuss & Levine, 2002). While the DL-PFC is part of the archicortical trend originating in the hippocampus, the OFC (which overlaps with the ventral prefrontal cortex) is part of the paleocortical trend emerging from the caudal orbitofrontal (olfactory) cortex. The DL-PFC consists of the cytoarchitectonic areas Broadmann’s areas (BA) 9, 46 and possibly 8. The OFC (also referred to as the ventromedial cortex) consists of BA 10, 11, 12, 25, and 47 (see Figure 1.2).

Functional distinctions between these two regions can also be made. In particular, the DL-PFC is critical in making decisions that call for the integration of multiple sources of information (Krawczyk, 2002). Damage to DL-PFC is associated
with typical executive function deficits, such as perseveration, planning difficulties and impulsivity (Stuss & Bensen, 1984). Alternatively, the OFC is involved in emotional processing, including the acquisition and reversal of stimulus-reward associations (Rolls, 2000). Damage to the OFC has a profound effect on human behaviour, including deficits in affect, emotional states, social abilities, deciding, and reasoning (Krawczyk, 2002).

The proposed links between prefrontal regions and executive functioning are based on three lines of evidence. First, much research has demonstrated deficits in executive skills following damage to the prefrontal cortex. For example, patients with damage to the prefrontal cortex have been found to exhibit impaired judgment, planning and decision-making (Stuss & Benson, 1986), poor temporal organization of behaviour (Fuster, 1997), reduced working memory capacity (Goldman-Rakic, 1998), and impaired response inhibition and task combination (Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999).

Second, functional neuroimaging studies indicate significant activation within the prefrontal cortex in unimpaired individuals performing executive function tasks. For example, based on their review of 275 studies using position emission tomography (PET) and functional magnetic resonance imaging techniques (fMRI), Cabeza and Nyberg (2000) concluded that prefrontal areas (BA 6, 44, 9 and 46) were involved in almost all high-level cognitive tasks, with particular prominence during working memory and memory retrieval tasks.

Third, lesion studies of non-human animals with damage to the prefrontal regions show significant impairment in tasks associated with executive function. For example, the studies by Diamond and colleagues (e.g., Diamond & Goldman-Rakic, 1989; Diamond, Zola-Morgan & Squire, 1989) demonstrated that only lesions of the prefrontal cortex in adult rhesus monkeys produced impaired performance on two
measures of executive control, the A-not-B task and delayed response tasks. However, it must be remembered that the prefrontal cortex is a richly interconnected system, with numerous connections and feedback loops between it and virtually all other parts of the brain (Stuss & Benson, 1986). Damage to, or loss of function, at any level of these connections may therefore produce similar deficits to those caused by lesions to the prefrontal cortex (Mayes & Daum, 1997). Thus, although the frontal lobe may play an integral role in the execution of executive functions, the integrity of the entire brain region may be necessary for efficient task performance (Anderson, 2002).

While there is consensus among researchers about the general nature of executive functions and the involvement of prefrontal regions of the brain, different opinions emerge when more precise descriptions of the construct are attempted and decisions about assessment instruments are required. A major issue is whether executive function should be treated as a unitary construct or whether it should be subdivided into components (Burgess, Alderman, Evans, Emslie & Wilson, 1998). If the latter position is accepted, then three further questions are raised. The first concerns the basis of the subdivision. Should it reflect the involvement of specific cognitive processes or the involvement of specific regions within the prefrontal cortex, or some other criterion? The second question pertains to the extent to which executive function can be fractionated without compromising its construct validity. Researchers’ views on these issues inform their decisions about assessment instruments, as discussed further below. The third question refers to the relations between the individual components of executive function.

Current Approaches to Executive Functions in Adults

A range of theoretical approaches and associated assessment techniques have been used in the study of executive functions in adults. The neuropsychological approach focuses on the consequences of damage to the frontal lobes on the
performance of a range of global tests, and often emphasises the unitary nature of executive functions. Other researchers proposed varying degrees of fractionation by suggesting that the executive system can be subdivided by assessing more specific components such as working memory or inhibition. Finally, multiple process models focus on the links between the core components of executive function. The purpose of the following sections is not to provide an exhaustive review of the executive function research involving adults, but rather to describe these three main approaches and provide some illustrative examples. This will serve as useful background to the developmental research in later chapters.

Neuropsychological Approach to the Assessment of Executive Functions

Traditionally, executive functions were conceptualized as a general overarching construct that included all supervisory or self-regulatory functions associated with the prefrontal regions of the brain (Gioia, Isquith, Kenworthy & Barton, 2002). Consistent with this conceptualization, global measures of executive functioning such as the Wisconsin Card Sorting Task and Tower of Hanoi/London were adopted. Gambling tasks are discussed in this section as they have also been used in the examination of performance of patients with damage to the frontal lobes of the brain, and successful performance requires the integration of a range of cognitive skills.

Wisconsin Card Sorting task. The most common global measure of executive function, the Wisconsin Card Sorting Task (WCST) was originally developed by Grant and Berg (1948). In this task, participants sort a deck of cards according to various dimensions. Four reference cards varying in colour, shape and number are placed in front of a participant (see Figure 1.3). Participants are then given a deck of response cards, each containing one to four identical figures (triangles, stars, crosses or circles) in one of four colours. The participants’ task is to match the cards in the deck with one of the four reference cards. They are not informed of the sorting principle. The
experimenter indicates whether or not they have sorted each card correctly, and they must use this feedback to discover the sorting rule. After the respondent has categorized 10 consecutive cards correctly, the sorting principle is changed without warning. The experimenter gives negative feedback to responses based on the previously correct sorting criterion. Several dependent variables are available including number of categories completed, the number of errors, perseverative responses, and the number of times the respondent made five correct responses in a row but failed to sort the 10 cards necessary to complete a category.

Reference Cards

(a) (b) (c) (d)

= red
= green
= yellow
= blue

Response Cards

Figure 1.3. Material used in the Wisconsin Card Sorting Task (modeled on Dehaene & Changeux, 1991). Correct sorting of the response card by the colour criterion would be location (a), correct sorting by number would be location (b), or correct sorting by shape would be location (c).

The WCST has demonstrated utility in discriminating adults with frontal lobe lesions from those with non-frontal injuries. For example, in her classic study, Milner
(1963) found that patients who had undergone prefrontal lobectomies for treatment of epilepsy showed a severe and consistent impairment on the WCST. Milner compared pre- and post-operative performances in patients with a dorsolateral frontal lobectomy, and found a significant increase in perseveration post surgery. An error is labeled as perseverative if patients continue to use a rule that was previously correct, even after negative feedback is provided, that is, they fail to shift from one sorting rule to another (Dehaene & Changeux, 1991). In contrast to dorsolateral frontal patients, Milner observed no excessive perseveration in patients who had undergone temporal, bilateral, hippocampal, or inferior frontal lobectomies.

Subsequent studies also found that groups of patients with frontal lobe damage tended to perform worse than groups of patients with non-frontal damage, but these investigations also documented substantial variability in WCST performance across subjects with frontal lobe lesions. For example, Drewe (1974) reported that medial frontal lesions may be even more detrimental to WCST performance than dorsolateral frontal lesions. In addition, this study indicated that left frontal lesions generally yield more total errors, though not necessarily more perseverations, than right frontal lesions (Drewe, 1974). Stuss et al. (2000) also found support for the sensitivity of the WCST to focal frontal brain damage, in comparison to nonfrontal regions. The authors qualified their findings, however, by suggesting that the sensitivity of the WCST depends on exactly which dependent measure is used and adherence to the test administration procedures (Stuss et al., 2000).

A recent meta-analysis of 24 studies (published from 1963 to 2001) examined the performance of patients with brain damage on the WCST (Demarkis, 2003). The findings indicated that patients with frontal damage achieved significantly fewer WCST categories and generate more perseverative errors than patients with nonfrontal damage. In particular, a large effect size was found for DL-PFC damage. Consistent with this,
Lombardi et al. (1999) suggested that the functional integrity of the right dorsolateral frontal subcortical circuits was critical for WCST performance.

Tower of Hanoi and Tower of London. The Tower of Hanoi (TOH) developed by Borys, Spitz and Dorans (1982) is also a highly recognized global measure of executive functioning. This ring transfer task utilizes an apparatus with three pegs and disks of varying sizes (see Figure 1.4). Participants are required to plan a sequence of moves that transforms an initial configuration (tower of disks on left peg) to a goal state (tower of disks on right peg). The movement of the disks is governed by rules such that larger disks cannot be placed on top of smaller disks, and that only one disk can be moved at a time. Performance on the task requires the formulation of sub-goals, and determining the most advantageous order of moves to obtain each of the sub-goals as well as the final goal.

![Figure 1.4. Apparatus used in the Tower of Hanoi Task.](image)

A variant of the TOH, the Tower of London (TOL), was specifically developed to examine the cognitive abilities of patients with frontal lobe lesions (Shallice, 1982). The TOL has similar task requirements, but differs from the TOH in that while all disks are equal in size, the three pegs differ in length, thereby limiting the number of disks
Development of Executive Functions

that can be held by each peg. For both tower tasks, a planning efficiency score derived from the number of moves required to complete the problem is used as the measure of executive functioning.

There is accumulating evidence that both the TOH and the TOL are sensitive to frontal lobe damage. For example, Goel and Grafman (1995) tested 20 adult patients with lesions in the prefrontal cortex on the TOH task. Performance was measured by a composite score based on accuracy, completion time, and problem difficulty. The results indicated that the performance of the frontal patients was significantly worse than that of controls. Further, both frontal patients and controls appeared to use the same general strategy to solve the problem, and both groups demonstrated the ability to plan and ‘look ahead’. The performance of frontal patients was explained in terms of an inability to see or resolve a goal-subgoal conflict. In addition, Anderson, Damasio, Tranel and Damasio (2000) described the long-term behavioural and cognitive results of damage to prefrontal cortex in two young adult patients who had sustained their injury prior to 16 months of age. The authors reported severely impaired performance on the TOH by both patients. In addition, although both participants were able to complete six categories on the WCST, one subject performed in the defective range on perseverative errors (33 errors), whereas the second patient committed an average number of perseverative errors (only 8).

Gambling task. The WCST and the TOH have traditionally been thought to involve the functioning of the DL-PFC. In recent years, however, attention has turned to the contribution of the orbital, ventromedial regions to the performance of executive tasks (Rogers et al., 1999). For example, Bechara, Damasio, Damasio and Anderson (1994) developed the Iowa Gambling Task and found that it was sensitive to damage of the OFC. In this task, participants are required to make a series of card selections that result in winning and losing money. They were given a stake of $2000 in play money
and asked to win as much money as possible by choosing cards from any of four decks. The four card decks are characterized by different reward-punishment profiles, such that decks A and B offer high rewards but higher penalties, resulting in overall loss, whereas decks C and D offer smaller rewards but minimal penalties, resulting in overall profit. Bechara et al. (1994) found that patients with OFC lesions are guided by immediate prospects rather than long-term consequences. That is, while unimpaired participants quickly learned as the procedure progressed to select cards from the advantageous decks, the OFC patients continued to select cards from disadvantageous decks. Even when they could describe the contingencies governing the task, these patients could not use this knowledge to guide their behaviour (Bechara et al., 1994).

**Evaluation of the Neuropsychological Approach**

The main appeal of adopting the neuropsychological approach and the use of global measures lies in the fact that by definition, executive function is a broad construct. It therefore seems appropriate that measurement of this construct should utilize tasks that may capture many aspects of executive functioning. These global tasks have good criterion validity as they have not only been relatively successful in discriminating people with frontal lobe injury from unimpaired people, but might also differentiate between patients with injuries corresponding to DL-PFC versus OFC lesion sites.

The construct validity of these executive function tests, however, can be questioned, as the precise nature of the cognitive processes involved in the performance of these tasks is difficult to determine (Phillips, 1997). Labeled the ‘task impurity’ problem, these global tasks incorporate a range of lower-order skills, such as expressive and receptive language, fine motor skills, visual perception, reading ability, and processing speed (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001). They are impure in the sense that they may make demands on a variety of other non-executive
cognitive processes that are supported by other brain structures outside of the frontal cortex (Rabbitt, 1997). For example, performance on the TOH demands a variety of intact functions, from planning and temporal ordering of potential moves, to spatial and visual imagery skills, to holding information in working memory. Similarly, successful performance on the WCST requires attribute identification, categorization, working memory, inhibition, selective attention, and utilization of verbal feedback provided in the context of a social interaction. Impairment of any of these skills may result in a poor overall score on the task in the absence of pure executive function deficits. Thus, a low score on a global test does not necessarily mean inefficient or impaired executive functioning (Ozonoff, Rogers, Farnham, & Pennington, 1993). Consequently, it may be difficult to determine the precise difficulty experienced by an individual who performs poorly on a global task (Ozonoff, 1997).

A further difficulty is that global executive tasks tend to have low reliability (Miyake, Emerson & Friedman, 2000). By definition, tasks can only be valid tests of executive function when they are novel. With practice, tasks become automatised and consequently may be performed in qualitatively different ways (Rabbitt & Lowe, 2000). Repeated encounters with these tasks therefore reduce their effectiveness in identifying the target executive functions. Additionally, people may adopt different strategies in different testing sessions (Miyake et al., 2000). Thus, the test-retest reliability of global tasks has been disappointingly low (Rabbitt, 1997). Even parallel forms of the tests are problematic, as while new content may be used, the task format itself remains the same (Phillips, 1997).

One remedy to overcome the problems associated with global tasks is to adopt a narrower conceptualization of executive functions, with tests that are more tightly constrained. As can be seen in the following sections, such approaches emphasise a
more basic process, attributing variations in executive function performance to limitations in a primary cognitive capacity such as working memory or inhibition.

*Working Memory Approach*

One of the most prominent frameworks that focused on a more specific executive process is Baddeley’s (1986) multi-component model of working memory. This approach equates executive functions with working memory, which is defined as a system for the temporary maintenance and manipulation of information (Baddeley, Della Sala, Gray, Papagno & Spinnler, 1997). The original model developed by Baddeley (1986) included three components: a) the phonological loop that maintains verbal information; (b) the visuo-spatial sketchpad that stores visual and spatial information; and (c) the central executive that integrates, co-ordinates and monitors information. A fourth component, the episodic buffer, was recently introduced into the model, and is proposed as a limited-capacity temporary storage system that integrates information from numerous sources (Baddeley, 2000). The central executive is most relevant to executive function but is not well understood. Baddeley (1996) suggested that the Supervisory Attentional System (SAS) proposed by Norman and Shallice (1986) is a good representative model for the central executive component of his theory. The SAS is hypothesized to be necessary for responding appropriately to novel situations and forming plans. In the absence of this system, behaviour will be governed by either action schemas that are triggered by environmental features that lead to distractibility and a loss of behavioural control, or by current activity that results in response perseveration and a difficulty in shifting mental set (Hughes, Russell & Robbins, 1994). The central executive performs at least four different functions (Baddeley, 1996). These include the coordination of separate task performances (as measured by dual-task performance), generating retrieval strategies (random number generation, fluency tasks), selectively attending to a particular stimulus while
simultaneously inhibiting a separate stimulus (task-switching), and the ability to hold and manipulate information from long-term memory (activation of long-term memory).

The dual-task paradigm is a frequently used to investigate the central executive. D’Esposito et al. (1995) conducted an fMRI study that explored coordination of two simultaneous tasks. Participants were administered two individual non-working memory tasks, a semantic judgment task and a spatial rotation task, which had previously been found to activate posterior regions of the brain. Cerebral activity associated with tasks performed in isolation was compared to cerebral activity when the two tasks were performed simultaneously. While frontal activation was not observed during the single-task conditions, activation of the prefrontal cortex was observed when both tasks were performed concurrently. The authors concluded that the DL-PFC is involved in the working memory systems recruited during dual-task performance.

A similar model of working memory has been proposed by Just and Carpenter (1992). According to the authors, this model approximates the part of the central executive in Baddeley’s (1986) model that deals with language comprehension. Working memory capacity involves the simultaneous storage and processing functions, which draw on a common pool of resources. Thus when task demands for resources are high (due to either storage or processing demands), then an individual’s overall working memory capacity may be compromised as the limited available resources cannot serve both processes simultaneously. The capacity of working memory may therefore be assessed by complex span tasks that involve simultaneous storage and processing demands (Daneman & Merikle, 1996). For example, Daneman and Carpenter’s (1980) Reading Span task required participants to read a set of unrelated sentences, and then at the end of the set, to recall the final word of each sentence. The task involved the presentation of increasingly larger sets of sentences, with the performance of college students varying from recall of two to five final words.
Engle (2002) proposed that working memory comprises both short-term memory and controlled attention components. According to this approach, keeping relevant information easily accessible depends on the individual’s ability to control attention. Thus, an individual’s working memory capacity may not necessarily reflect their ability to store information but rather the ability to maximize retention of information through the suppression of irrelevant stimuli or responses. In support of this view, Kane, Blecky, Conway and Engle (2001) examined the performance of participants with high and low working memory span on a prosaccade task, in which a visual cue appeared in the same location as a subsequent to-be-identified target letter, and on an antisaccade task, in which a target appeared opposite the cued location. While both span groups performed comparably on the prosaccade task, the low-span participants had significantly longer mean target identification times and poorer accuracy for the antisaccade task than their high span counterparts, reflecting differences in attentional control.

Evidence for the association between working memory and the prefrontal cortex was presented by Cabeza and Nyberg (2000) who, as noted earlier, reviewed 275 PET and fMRI studies of a variety of cognitive tasks. The results indicated that the prefrontal cortex was involved in almost all high-level cognitive tasks, however, prefrontal activations (BA 6, 44, 9, and 46) were particularly prominent during working memory and memory retrieval.

Further research has suggested that the subcomponents of the prefrontal cortex might be functionally distinct. For example, Goldman-Rakic (1987) argued for a functional segregation based on the type of information being maintained, with the DL-PFC principally involved with spatial information and the ventro-lateral PFC (VL-PFC) with nonspatial information (Goldman-Rakic, 1987).

A hierarchical two-stage model of the involvement of the prefrontal cortex in working memory was proposed by Owen, Evans and Petrides (1996). Within this
Development of Executive Functions

framework, the VL-PFC is proposed as the site where information is initially received from the posterior cortex, with some theorists also suggesting that the VL-PFC guides active online maintenance of these accessed representations (D’Espositio et al., 1998). The DL-PFC is then proposed to be recruited when additional monitoring and manipulation of information in working memory is required. The VL-PFC and DL-PFC are therefore hypothesized to be hierarchically related, with dorsolateral regions operating on the products of the VL-PFC.

Owen et al. (1996) used position emission tomography with magnetic resonance imaging to study performance on five spatial tasks that varied in terms of their executive function demands. Tasks requiring the organization and execution of a sequence of spatial moves retained in working memory produced significant changes in blood flow to the ventrolateral frontal cortex (BA 47) bi-laterally. However, tasks requiring the active monitoring and manipulation of spatial information within working memory produced additional activation in mid-dorsolateral frontal cortex (BA 46 and 9), thereby supporting the two-stage model of spatial working memory.

Several other studies have also supported this VL-PFC/DL-PFC division in working memory. For example, Rowe, Toni, Josephs, Frackowiak, and Passingham (2000) used event-related fMRI to study the performance of humans on a spatial working memory task. The authors distinguished between the maintenance of spatial items and the selection of an item from memory to guide a response. The results indicated that maintenance of spatial representations was associated with marginal activation of the DL-PFC, whereas selecting between different representations produced significant DL-PFC activity.

Wagner, Maril, Bjork, and Schacter (2001) also used an event related fMRI study to investigate the distinct subregions within the prefrontal cortex during the performance of verbal working memory tasks. Participants performed two executive
control tasks that differed in the types of control processes required; a rote rehearsal task that required phonological access and maintenance, and an elaborative rehearsal task that required both maintenance and monitoring of the task items. While rote rehearsal preferentially activated VL-PFC, elaborative rehearsal preferentially activated DL-PFC. Additionally, a significant temporal difference was identified between the two regions. The DL-PFC response systematically lagged behind the VL-PFC response, thereby supporting a hierarchical model.

_Evaluation of the working memory approach._

Working memory approaches have not only accounted for individual differences in the performance of some executive function tasks, but have also allowed for the differentiation of separable brain regions. However, working memory approaches may not provide a complete account of executive functioning. For example, Lehto (1996) examined the relations between working memory capacity (including digit and word span, a modified memory-updating task, and five different complex span tasks) and three executive function tests (WCST, TOH and Goal Search Task). Whereas performance on the WCST correlated significantly with working memory tasks, the TOH and Goal Search task did not. Thus, the working memory approach may have difficulty accounting for executive function tasks that have low working memory demands.

_Inhibitory Approach_

A second executive process that has been examined is that of inhibition, which has been defined by Dempster and Corkill (1999) as the ability to suppress irrelevant information during the execution of a plan. Aron, Robbins and Poldrack (2004) identified three aspects of inhibition: 1) response inhibition, or the cognitive process required to cancel an intended movement; 2) task-switching which involves the executive control required to change from performing one task to another; and 3)
inhibition during memory retrieval which entails the control of distracting memories. Each of these processes can be assessed independently from each other.

Response inhibition can be examined using the go/no-go task. In this task, participants are generally required to display a simple motor response to one cue (the go stimulus), while refraining from responding to another stimulus (the no-go stimulus). Scoring is based on reaction time, errors of commission (incorrectly responding to a no-go stimulus), and errors of omission (not responding to a go stimulus). A number of studies have reported that patients with frontal pathology are impaired on the go/no-go task (Decary & Richer, 1995; Leimkuhler & Mesulam, 1985). Liddle, Kiehl and Smith (2001) used event-related fMRI to measure the performance of 16 healthy subjects during a go/no-go task. The results indicated that the anterior cingulate cortex was active during both the go and no-go trials, DL-PFC and VL-PFC was more active during the no-go trials, while primary motor cortex, supplementary motor area, pre-motor cortex and cerebellum were more active during go trials. The authors concluded that while the anterior cingulate cortex is principally involved in making and monitoring of decisions, the DL-PFC and VL-PFC regions are specifically engaged during response inhibition.

The second inhibitory process identified by Aron et al. (2004) was that of task-switching. The task-switching paradigm requires participants to change from performing one task to another. The dependent variable is switch cost, which measures the time taken to switch between tasks compared with the time taken to repeat a task. Rogers and Monsell (1995) utilized this paradigm to examine the performance of adults who were required to switch between two tasks, a letter task and a digit task. For both tasks, participants were shown a compound stimulus of two characters, such as ‘N2’. Participants classified the letter as a consonant or a vowel in the letter task. They classified the digit as even or odd in the digit task. The result provided evidence of
significant switch costs as the participants were much slower to respond when they were required to switch from one task to the other. Evidence exists for the involvement of the prefrontal cortex during task switching performance. For example, DiGirolamo et al. (2001) used fMRI neuroimaging to examine the performance of adult participants during task switching. The results indicated significant activation of the medial and dorsolateral frontal regions during the switching trials.

Paired associate tasks are used to assess inhibition during memory retrieval (Aron et al., 2004). In the AB-AC version of this task, participants learn two lists of word pairs, with each pair consisting of cue words (e.g., bird) and response words (e.g., bath). The participant is first required to recall the response words when presented with the cue words. However, a second list of paired words is then presented so that the same cue words (bird) are paired with new response words (e.g., dawn). Learning and recall of the second response list is vulnerable to interference from the first list. Frontal lobe functioning has been implicated in the performance of paired-associates tasks. For example, patients with damage to the DL-PFC showed greater vulnerability to interference in the AB-AC paired-associates task (Shimamura, Jurics, Mangels, Gershberg and Knight, 1995).

Inhibitory control processes of patients with frontal lobe damage have also been investigated using the Stroop Colour-Word Test (Stroop, 1935). The Stroop task requires a participant to name the colour of the ink in which an incongruent word is written (e.g., the word ‘red’ printed in green ink). The time taken to name the ink colour is usually compared to some other timed measure (e.g., naming colours on a colour chart) to provide an estimate of interference. To perform this test successfully, participants must pay attention to the relevant stimulus dimension (ink colour) and inhibit the irrelevant but prepotent process of reading and naming the printed word. Current available evidence suggests that the Stroop Test tends to be disproportionately
difficult for subjects with frontal lobe damage. For example, Stuss, Floden, Alexander, Levine and Katz (2001) examined the performance of patients with single focal brain lesions in frontal and non-frontal lesions. Only those with frontal lesions were significantly impaired on the Stroop. More specifically, damage to the left dorsolateral frontal lobe resulted in increased errors and slowness in response speed when naming the colours on a colour chart, while bilateral superior medial frontal damage was associated with increased errors and slowness in response times in the incongruent condition.

In addition, Bench et al. (1993) used PET to measure changes in regional cerebral blood flow during the performance of normal subjects on the Stroop task. Results indicated involvement of the anterior right hemisphere and the medial frontal structures during the performance of this task.

Evaluation of the inhibitory approach.

As with the working memory approach, focusing on a specific aspect of executive functioning like inhibition may compromise construct validity. Explanations that accentuate a specific central process, such as working memory or inhibition, may be too simple to provide an adequate characterization of the way in which different aspects of executive function interact to accomplish complex metacognitive processes (Zelazo & Müller, 2002; Stuss & Benson, 1986). The involvement of similar prefrontal regions (e.g., DL-PFC) in a large number of executive function tasks (including the WCST, TOH and delayed response task) also suggests that executive functioning may be best understood in terms of interactions between a number of different functions, rather than in terms of a specific association between one region and a single higher-level cognitive process (Collette & Van der Linden, 2002).
Multiple Component Models

In light of these criticisms, investigators have also begun to examine the links that may exist between these component processes. Such studies conceptualise executive functions as multiple process related systems that are interrelated and interdependent. For example, Roberts and Pennington (1996) suggest that inhibition and working memory are most often interactive, with working memory load influencing the extent of inhibitory responses. The authors cite evidence from an earlier study that demonstrated that manipulating working memory load dramatically alters the ease of inhibition, even in normal adults. In this study, Roberts, Hager and Heron (1994) tested the competing resource hypothesis between inhibition and working memory using the Antisaccade task, described above. Participants made incorrect reflexive saccades 25-30% of the time, however, performing a simple single-digit arithmetic task at the same time increased reflexive saccading to 51%. In addition, two other secondary tasks that consisted of the same input-output requirements as the arithmetic secondary task, but that had a lower working memory load, did not produce increased prepotent responding in the antisaccade task. Thus, inhibitory control was more difficult when working memory load was higher.

The links between the various component processes involved in executive functioning were investigated by Miyake, Friedman, Emerson, Witzki and Howerter (2000). This study used latent variable analysis to determine to what extent different executive functions can be considered unitary (i.e., reflect the same underlying mechanism) or non-unitary. A set of tasks postulated to reflect three executive functions (mental set shifting, information updating and monitoring, and inhibition of prepotent response) was administered to 137 college students. Confirmatory factor analysis indicated that these three executive functions are moderately correlated with one another, however, they were also clearly distinguishable as the full three-factor model
produced the best fit of the data. Further, structural equation modeling suggested that these three executive functions contribute differentially to performance on global executive tasks. Performance on the WCST was related most strongly to the shifting function while performance on the TOH was more clearly related to inhibitory process. The authors concluded that while the various executive functions can be distinguished from each other, they nonetheless remain related in that they share some underlying commonality.

_Standardised Test Batteries._ A number of standardized test batteries of executive functioning have been recently developed. The Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan & Kramer, 2001) for example, provides a standardized assessment of executive function in children, adolescents, and adults between the ages of 8 and 89. It is composed of nine stand-alone tests that can be individually or group administered and include: Trail Making, Verbal Fluency, Design Fluency, Colour-Word Interference, Sorting, Twenty Questions, Word Context, Tower, and Proverb. The D-KEFS was co-normed on a large and representative national sample and is designed exclusively for the assessment of executive functions.

The Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie, & Evans, 1996) has also been used extensively in research investigating adult executive dysfunction. This test battery includes six subtests including a Rule Shift Card, Action Program, Key Search, Temporal Judgment, Zoo Map and Modified Six Elements task. The BADS also includes a 20-item questionnaire about everyday executive problems, known as the Dysexecutive Questionnaire, which has been used effectively as an individual screening instrument to identify executive dysfunction (Bennet, Ong & Ponsford, 2005).

A criticism of these standardized test batteries, however, lies in the lack of theoretical foundations underpinning their construction and use. The subtests of the
BADS, for example, were selected purely on the basis of their previous individual use in identifying executive dysfunction in impaired samples. While this supports the discriminate validity of standardized test batteries at a broad level, such approaches might contribute little to the continuing questions regarding the nature of executive function as a construct. In the absence of evidence identifying the cognitive processes in the subtests and a theory about how such processes relate to one another, there can be no principled means to interpret either the individual subtest scores or the global scores. In most cases, there is no attempt to consider whether the various subtests are equivalent in terms of task complexity. This further complicates interpretation of test results and dissociations between clinical and control groups. For example, if subtest A (proposed to measure planning) is more complex that subtest B (thought to measure attention) and it is found that a clinical group is impaired relative to the control group on task A, but not task B, it would be inappropriate to conclude that the clinical group had poorer planning capability but comparable attention. The observed pattern might instead be due to the clinical group’s difficulty in coping with the greater complexity of task A. Conclusions regarding differential impairments may therefore be difficult if tasks differ in complexity as well as the cognitive processes they are intended to assess.

Conclusions

While the involvement of frontal lobes in the performance of executive function tasks has largely been established, questions remain regarding the operationalisation of the construct and associated assessment procedures. To date, the assessment of executive functions in adults has largely been guided by approaches that focus on neuropsychology, working memory, inhibitory control, multiple component and standardized test batteries. While each of these approaches has offered unique information to the research literature, each on its own may be an incomplete account of the complex cognitive processes that fall under the umbrella of executive functions. To
a large extent, however, these approaches have guided research on the assessment of executive processes in childhood, as will be outlined in the next chapter.
CHAPTER 2. DEVELOPMENT OF EXECUTIVE FUNCTIONS

More recently, the focus on executive functions has turned towards the development of these cognitive processes during childhood. Impairments in executive function have been implicated in a wide variety of developmental and psychological disorders including autism (Griffiths, Pennington, Wehner, & Rogers, 1999), ADHD (Barkley, 1997), phenylketonuria (Griffiths, Tarrini & Robinson, 1997), Turner’s syndrome (Temple, Carney & Mullarkey, 1996), Tourette Syndrome (Gedye, 1991), conduct disorder (Pennington & Bennetto, 1993), and obsessive-compulsive disorder (Enright & Beech, 1993). Further research is required, however, to assess if these different developmental disorders might have specific patterns of strengths and weaknesses that may allow them to be differentiated on the basis of their unique executive profiles (Ozonoff, 1997; Ozonoff & Jensen, 1999). Thus, the continued assessment of executive functions in both normally developing and clinical child samples is critical.

It has been suggested that functional improvements on executive function tests parallel the biological development of the prefrontal cortex (Anderson, 2002). The biological development of the frontal lobes will therefore now be outlined.

Biological Development of the Frontal Lobes

Researchers investigating the biological maturation of the brain have demonstrated that morphological development in the frontal cortex is incomplete at birth (Stuss, 1992), and that the frontal lobes are in fact the last area of the brain to fully develop (Dempster, 1992). From birth to the second year of life, the area of the frontal lobes increases sharply in size. This is followed by a second less pronounced growth spurt between four to seven years of age. A slow, but much less dramatic increase in the size of the frontal lobes then continues until young adulthood (Dempster, 1992). While much of this growth has been attributed to an increase in the size and complexity of
nerve cells, other significant changes have also been noted. For example, while myelination of the sensory and motor areas is almost complete by two years of age, pathways of the frontal lobes are among the last of all brain areas to fully myelinate, with this process continuing up to about age twenty. Similarly, developmental changes in synaptic density have also been reported throughout childhood. According to Huttenlocher (1984; 1999), synaptic density increases during infancy reaching a peak at one to two years of age, with the number of synapses in infancy exceeding adult numbers by approximately 50%. Following this exuberant period of synapse accumulation, excess synapses are progressively eliminated throughout childhood and adolescence (Goldman-Rakic, 1987; Huttenlocher, 1979). As the process of synapse elimination involves the removal of unproductive neural tissue, it has been suggested that adult levels of competence may depend upon an optimal level of synaptic pruning (Goldman-Rakic, Bourgeois & Rakic, 1997).

Furthermore, studies using resting EEG recordings in normal children have reported evidence for stages of neural maturation that appear to coincide with qualitative changes in cognitive development. For example, Thatcher (1991, 1992, 1997) identified a cycle of brain electrical signal development between one and five years of age. These cycles of EEG activity are interpreted as repetitive sequences in which periods of overproduction of synapses were followed by a phase of pruning. Thatcher (1997) models this process as a developmental spiral staircase in which brain structures are periodically revisited and refined, resulting in step-wise increases in cortical connections and interconnections within the frontal lobes. Thus, the progression of frontal lobe maturation is not a smooth course, but is punctuated by growth spurts and separable phases of development (Grattan & Eslinger, 1991). Additionally, different areas of the frontal lobes may follow different maturational timetables. For example,
Zelazo and Müller (2002) stated that the OFC typically develops earlier than the DL-PFC and comes to be regulated by it.

It seems clear that the structure and function of the frontal cortex changes significantly in the preschool-school period, including large-scale pruning of synaptic connections and myelination of subcortical prefrontal pathways. Associated with this ongoing maturation, children and adolescents gradually acquire the capacity for more efficient cognitive processing. This suggestion has been supported by a range of developmental researchers who have charted periods of rapid development in executive function performance during childhood (e.g., Hughes, 1998; Welsh, Pennington & Grossier, 1991). Measurement of developmental changes in executive functions during this period is therefore particularly important.

*Executive Function and Cognitive Development*

Executive functions emerge in the context of normal cognitive development, and the two are likely to be closely linked. Therefore it will be useful to outline some theories of cognitive development so that potential links with executive function can be identified. Executive functions are most often characterized as domain-general in their application. Therefore the focus here will be on domain-general theories of cognitive development, rather than those that apply only to specific content domains.

*Piagetian Stage Theory*

Classic developmental theories, such as Piaget (1954), adopted a domain-general approach to cognitive development that involved the transitions between stable, pervasive stages. Most summaries of Piaget’s theory focus on the four general structural stages that were identified in cognitive development, with each stage representing a distinctive organization or structure in children’s conceptual understanding. In the sensorimotor stage (0 to 2 years) infants construct an understanding of the world by coordinating sensory experiences with physical action. In the preoperational stage (2 to
7 years) children begin to symbolically represent the world with words, images and drawings. The concrete operational stage (7 to 11 years) involves the ability to perform operations that involve concrete objects and to reason about concrete examples. The final formal operational stage (11 to 15 years) sees individuals move away from concrete experiences and to think in more abstract and logical ways. Hence Piaget drew a sharp distinction between the thought of young children, who were limited to intellectual operations that could be applied directly to existing situations, and that of adolescents who were capable of integrating such operations within a single system and then testing the hypotheses that this system permitted them to generate (Marini & Case, 1994). Progression through these stages occurred as a result of the child’s search for equilibration, or to resolve the competitive contradictions or conflicts among schemes elicited by different situations. Thus, the interplay between perceiving the world as it is (accommodation) and simultaneously interpreting the world through currently available mental models and representations (assimilation) resulted in mental structures being built, revised and transformed to reach a balance, or equilibrium of thought.

**Neo-Piagetian Theories**

Neo-Piagetian theories of cognitive development, which include those proposed by Case (1992) and Pascual-Leone (1970), emphasise different mechanisms for cognitive growth. These theories include an executive function like component that accounts for growth in cognitive development. For example, Case’s theory (1992; Case & Okamoto, 1996) proposed that developmental change was driven by changes in working memory capacity. Case’s conception of working memory corresponds to Just and Carpenter’s (1992) working memory model whereby limited resources are available for simultaneous storage and processing. More specifically, it was suggested that experiences during childhood promote growth and elaboration of conceptual systems that leads to processes becoming more automatic. Thus, development is not simply a
matter of increases in resources. Rather, faster and more efficient processing makes available additional resources from a set capacity limit, which may then be allocated for the storage of information. Case identified at least three different conceptual domains that are aided by these developments in working memory: number, social cognition, and spatial relations. For example, with adequate experience, children’s level of numerical understanding was suggested to develop through four levels beginning at approximately age four, at which time they do not quantify correctly and ending at approximately age 10 when children are able to quantify along two dimensions simultaneously and relate those two domains precisely. Insufficient experience might lead to number knowledge skills at a level below what would be expected at a given age. However, efforts at enriching a child’s level of understanding would not necessarily lead to advanced knowledge because of limitations in working memory capacity.

Similarly, Pascual-Leone (1970) proposed a mathematical model that also included a capacity construct that was similar to working memory. According to this theory, mental attentional capacity is an innate resource that grows in power as children develop. Labeled the Theory of Constructive Operators, this model included two levels of psychological constructs: 1) schemes, which were derived from Piaget’s theory and represent the information-bearing, situation specific constructs that generate performance; and 2) hardware operators which were proposed as noninformational, content-free processing resources, such as mental attention and structural learning, that provide activation energy for the schemes. Cognitive performance is therefore determined by the interaction of schemes and hardware operators. Input leads to activation of a number of schemes which constitute the “field of mental attention” or working memory. The selection and activation of these schemes depends on three mechanisms: 1) the M-operator, a central activatory “mental energy” mechanism; 2) the I-operator, a central inhibitory attentional interrupt; and 3) the executive schemes,
which are responsible for the control of performance. The size of the M-operator is limited and increases as a function of chronological age. Empirical evidence suggests that M-capacity increases by one informational unit every second year, from one at the age of 3 to the adult capacity of seven at the age of 15 (e.g., Johnston, Fabian & Pascual-Leone, 1989). It is the growth of this M-capacity that is suggested to account for developmental change in cognitive performance, leading to greater capacity for coping with task complexity (Pascual-Leone, 2000).

In summary, both Case’s and Pascual-Leone’s approaches seek to explain the development of higher order processes (thinking, reasoning) and they include a capacity construct (working memory or M-capacity) which develops during childhood. That is, cognitive growth, including the development of executive functions, may be viewed as a consequence of increases in working memory capacity. Thus, these approaches can be likened to working memory approaches to executive function.

Complexity Approaches to Cognitive Development

More recently, research in cognitive development has emphasized the importance of complexity. Two recent approaches will be considered here, the Cognitive Complexity and Control (CCC) theory and Relational Complexity theory.

**CCC Theory.** According to CCC theory developed by Zelazo and colleagues, age-related changes in the complexity of the rule systems that children are able to use permit children to exercise increased control over their thoughts and actions, leading to qualitative changes in executive functioning (e.g., Frye, Zelazo & Palfai, 1995; Zelazo & Frye, 1998). CCC theory suggests that higher order rules make possible reflective awareness of lower order rules and make possible the deliberate selection of lower order rules. The theory postulates that children first acquire the ability to use a pair of arbitrary rules, typically around three years of age, and they then acquire the ability to integrate two incompatible pairs of rules into a single rule system via a higher order
rule, which usually occurs around five years (Zelazo, Jacques, Burack & Frye, 2002). These increases in control are said to have widespread consequences for children’s reasoning in social and nonsocial domains.

![Hierarchical tree structure depicting formal relation among rules (Zelazo & Frye, 1998). Note: s₁ and s₂ = setting conditions; a₁ and a₂ = antecedents; c₁ and c₂ = consequences.](image)

Figure 2.1. Hierarchical tree structure depicting formal relation among rules (Zelazo & Frye, 1998). Note: s₁ and s₂ = setting conditions; a₁ and a₂ = antecedents; c₁ and c₂ = consequences.

The tree structure in Figure 2.1 illustrates the way in which one rule can be embedded under another and can be controlled by it. If-then rules are expressed in terms of antecedents and consequents. As explained by Zelazo and Frye (1998), Rule A, which indicates Consequent 1 (c₁) should follow Antecedent 1 (a₁) is incompatible with Rule C, which connects a₁ to a different consequent (c₂). Rule A is embedded under, and controlled by, a higher order rule (Rule E) that can be used to select Rules A and B, as opposed to Rules C and D. In order to deliberately choose between Rules C and D, on the one hand, and Rules A and B, on the other, children need to be aware of the fact that they know both pairs of rules. Representation of a higher-order rule (E) that integrates the two pairs of rules (A and B; C and D) depends on the child’s ability to reflect on the lower order rules (Zelazo & Frye, 1998). It should be noted that CCC theory has recently been revised (CCC-r; Zelazo, Müller, Frye, & Marcovitch, 2003),
however, the foundations underlying the CCC-r are consistent with the original theory. The revision aims to incorporate the Levels of Consciousness model (Zelazo, 2000) and suggests that these age-related changes in maximum rule complexity are made possible by age-related increases in children’s highest level of consciousness (that is, the degree to which children can consciously reflect on the rules they represent).

CCC theory has been used to analyse the complexity of representational rule use tasks. The Dimensional Card Change Sort task (DCCS) is similar to the WCST in that children are required to sort cards according to their shape and colour. However, unlike the WCST, children do not have to discover the sorting rule. Children are shown two target cards (e.g., a red car and a blue flower) that remain visible throughout the task, and children are asked to sort test cards (e.g., blue cars and red flowers) that match one of the target cards on one dimension (e.g., shape) and match the other target card on the other dimension (e.g., colour). Children are first told explicitly to sort test cards by one dimension (e.g., colour) for a number of trials and then they are told to sort cards by the other dimension (e.g., shape). Regardless of which dimension is presented first, research indicates that 3-year-olds typically continue to sort the cards by that dimension, despite being told the relevant rules on each trial (Zelazo & Frye, 1997). Thus, while children can easily consider a pair of rules simultaneously (e.g., “If red …here; if blue …there”), it is not until five years of age that children can typically represent a higher order rule that allows them to select between two incompatible pairs of rules (e.g., “If we’re playing the colour game, then if red flower … there, and blue car … here, but if we are playing the shape game, then if red flower … here, and if blue car … there”). The ability to reflect on lower order rules through the formulation of higher order rules is therefore the process by which the development in responding in accordance with rules occurs.
Relational Complexity Theory. According to Relational Complexity theory, many higher cognitive processes, such as executive functions, involve the processing of relations. Relational complexity refers to the arity of relations, that is, the number of related entities or arguments inherent to a task or concept (Halford, Wilson & Phillips, 1998). Each argument corresponds to a dimension, and an $N$-ary relation is a set of points in $N$-dimensional space. The number of dimensions corresponds to the number of interacting variables that constrain responses or decisions. Relational processing is quantified in terms of a metric of relational complexity which defines unary relations as having a single argument as in class membership, dog(fido), a binary relation as having two arguments as in larger-than(elephant, mouse), ternary relations as having three arguments as in addition(2,3,5), quaternary relations as having four interacting components as in $2/3 = 6/9$, and quinary relations as having five interacting components as in $r(v,w,x,y,z)$ (Halford et al., 1998). As the relational complexity of a cognitive task increases, so too does its processing load.

Halford et al. (1998) also proposed two mechanisms whereby the complexity of concepts, and subsequently their processing loads, can be reduced. Firstly, conceptual chunking involves recoding concepts into less complex relations. For example, speed = distance/time is a ternary relation, but speed can be recoded into a unary relation, speed(60 kph), as when speed is indicated by the position of a pointer on a dial. Conceptual chunking reduces processing load, but leads to a temporary loss of access to relations between chunked variables. Secondly, complexity can be reduced through the segmentation of tasks into less complex steps, which can then be processed serially. Such serial strategies can be taught, or segmentation can be facilitated by experimental manipulations. Segmentation also reduces processing load, but relations between variables in separate chunks become inaccessible. Thus, both the strategies of chunking and segmentation can reduce processing load by reducing the number of dimensions.
that must be processed in parallel. However, chunking and segmentation strategies cannot be applied if relations between chunked variables or between variables in different segments must be used in making a decision (Halford et al., 1998). Thus, tasks that are experienced as difficult are often structured in a way that resists decomposition (e.g. Andrews & Halford, 2002).

Developmentally, the complexity of relations that can be processed by children has been shown to increase with age. Preliminary estimates were that unary relations can be processed by a median age of 1 year, binary relations by 2 years, ternary relations by 5 years, and quaternary relations by 11 years of age (Halford, 1993). Thus, from this perspective successful performance on executive function tasks would only be observed when the child had the appropriate level of processing capacity to effectively handle the complexity presented by the particular executive function task.

Relational complexity has been found to be an effective measure of functioning. For example, Andrews and Halford (2002) conducted complexity analyses of tasks in the domains of transitivity, hierarchical classification, class inclusion, cardinality, comprehension of relative clause sentences, and hypothesis testing and concluded that each was ternary-relational. They then designed items at a lower level of complexity (binary) for each domain. Thus, relational complexity was manipulated while other factors were tightly controlled. The tests at both levels of complexity were administered to 3- to 8-year-old children. The results indicated a strong effect of relational complexity, with the binary-relational items of each task being passed earlier than the ternary-relational items. Additionally, significant correlations were found between the items at the same level of complexity across the different task domains.

**CCC and RC theories.** In comparison to CCC theory, the relational complexity approach is based on the number of variables that are related rather than on the levels of the hierarchy of rules. For example, CCC theory suggests that the DCCS task entails a
Development of Executive Functions

rule for assigning colours, a rule for assigning shapes, and a higher-order rule that allows switching between shape and colour sorting. According to relational complexity theory, the higher-order rule comprises an additional dimension, thereby relating three variables: the setting condition (shape, colour), sorting-card attribute and sorting category. The typical DCCS problem is therefore ternary relational. Although CCC theory and the relational complexity are inter-translatable to some extent and make similar predictions in many situations, relational complexity differs in that it applies to both hierarchical and nonhierarchical processes (Halford et al., 1998). Relational complexity theory also includes principles that govern segmentation and chunking. In some situations, this might allow more precise predictions to be generated. For these reasons, complexity manipulations in the current research will be based on the relational complexity approach.

_Delineating Functions of the Prefrontal Cortex._

Recent investigations have suggested that the development of executive functions may not occur in a domain general fashion, and instead discrete regions of the prefrontal cortex may be dedicated to particular processes that may develop at different times (e.g., Zelazo & Müller, 2002). Specifically, a distinction might be possible between the ‘hot’ aspects of executive function such as the experience of reward and punishment or the interpretation of complex emotions that are associated with the OFC, and the more purely cognitive ‘cool’ aspects such as mechanistic planning, verbal reasoning or problem solving that are associated with the DL-PFC.

Two developmental approaches have been outlined. Firstly, Metcalfe and Mischel (1999) proposed a two-system framework for understanding the processes that enable self-control. The authors differentiated between a cool, cognitive “know” system that is specialized for complex spatiotemporal and episodic representation and thought,
and a hot, emotional ‘go’ system that is specialized for quick emotional processing and responding on the basis of unconditional or conditional trigger features.

Secondly, Zelazo and Müller (2002) distinguish between ‘cool’ executive functions that are associated with the DL-PFC and ‘hot’ executive functions associated with the OFC. Cool executive functions are more likely to be elicited by relatively abstract, decontextualised problems, for example, the WCST or the DCCS, as each require sorting by dimensions such as number and colour. In contrast, hot executive functions are required for problems that involve the regulation of affect and motivation. This may include tasks that involve the consideration of the affective significance of rewards, such as gambling tasks.

It was also suggested that the OFC develops earlier than the DL-PFC and comes to be regulated by it. As such, damage to the OFC would have cascading consequences for the development of the DL-PFC (Zelazo & Müller, 2002). For example, the hot/cool theory argues that important aspects of autism may be understood in terms of a primary deficit in the affective ‘hot’ aspects of executive function. This then has cascading consequences for the development of the DL-PFC, which leads to associated impairments in cool executive functions. Deficits in both the hot and cool domains would therefore be expected in autistic samples. On the other hand, important aspects of ADHD may be understood in terms of a primary deficit in the cognitive ‘cool’ aspects. According to the hierarchical relation between the OFC and the DL-PFC, damage to the DL-PFC will have relatively fewer consequences for the OFC, sparing many of the hot executive functions that are impaired in autism (Zelazo & Müller, 2002). Current research conclusions regarding the primacy of hot versus cool executive function impairments are limited, however, as they have not taken complexity into account. That is, tasks currently used in investigations of hot and cool executive functions might differ in terms of the complexity of the cognitive processes that the tasks require (Zelazo &
Müller, 2002). Therefore comparisons across tasks may be misleading because these tasks vary in terms of the demands they place on participants.

An example of the potential confound of complexity may be found in a recent study by Hongwanishkul, Happaney, Lee and Zelazo (2005) who administered two cool measures (Self-ordered pointing, DCCS) and two hot measures (Children’s Gambling Task and Delay of gratification) to a sample of 3- to 5-year-old children. Age-related improvements were found across all four executive function tasks. However, whereas the two cool measures were significantly positively correlated to one another, the two hot measures were negatively correlated (when age was partialled out). Additionally, while the cool measures were related to measures of general intelligence and temperament, the hot measures were not. While the hot and cool tasks showed different patterns of correlations with each other and the control measures, conclusions regarding the emergence of these abilities may be confounded by the complexity of the tasks involved. Future investigations would benefit from controlling for task complexity. For example, each task could also include items presented at a lower level of complexity in order to ensure that even the youngest children can understand the task instructions and procedures.

With respect to clinical samples, Dawson, Meltzoff, Osterling and Rinaldi (1998) examined the performance of children with autism, Down Syndrome, and controls on a delayed non-matching to sample task (associated with the medial temporal lobes and OFC) and a delayed response task (associated with the DL-PFC). Children with autism performed significantly worse than children with Down Syndrome and typically developing children on both tasks. However, for the autistic sample the severity of autistic symptoms was strongly associated with performance on the medial temporal lobe task and OFC, but not the DL-PFC task. While the results provides initial support for a primary deficit of hot executive functions in autism, the complexity of the
cognitive processes required by the two tasks used in this study was not examined. Therefore it remains unclear whether the differential impairments demonstrated by the two clinical groups were due to the cognitive processes required by the individual tasks, or the level of complexity required to succeed.

One domain of hot executive function to which complexity theories have been applied is theory-of-mind. Theory-of-mind is considered a hot executive domain as, like other hot tasks, it requires flexible problem representation in which children are required to consider problems from multiple viewpoints (Zelazo & Müller, 2002), and it has been linked to the OFC through imaging studies (Frith & Frith, 1999). A complexity account of theory-of-mind development has been applied by Frye et al. (1995) who attempted to account for children’s performance on theory-of-mind tasks in terms of improvements in reasoning based on embedded rules. Use of embedded rules was said to enable children to switch between judgments that may be based on two conflicting perspectives. According to Frye et al. (1995), setting conditions enable children to accomplish this switching, by appropriately selecting or restricting the application of a rule. The results of correlational analyses were interpreted as indicating that advances in performance on theory-of-mind tasks, the DCCS task, and a physical causality task between the age of three and five depend on being able to switch judgments across conditions, and that reasoning by embedded rules could account for these changes. However, in Experiment 3 only the DCCS version involving sorting by 1 of 2 dimensions was significant in predicting theory-of-mind performance. When the effects of age were partialled out, the card-sorting task accounted for 10% of variance in false belief and 6% of the variance in representational change theory-of-mind tasks (Frye et al., 1995).

Relational complexity theory has also been applied to the performance of theory-of-mind tasks. Andrews, Halford, Bunch, Bowden and Jones (Exp. 1 & 2, 2003)
examined the performance of normally developing children on standard theory-of-mind tasks (appearance reality and false belief items) as well as on some tasks from the cool content domains (transitivity, cardinality, and class inclusion, hierarchical classification) that involve binary and ternary items. In order to respond correctly on a typical theory-of-mind task (e.g., an appearance-reality task) children must appreciate that the relation between an object, and a person’s perception of that object, is conditional upon a third variable, the viewing condition under which the object is presented. Thus, the appearance reality tasks are equivalent to a ternary relation (setting condition, object attribute, object perception), and therefore children should have difficulty with this task until a median age of five years. Andrews et al. (2003) found that performance on relational complexity tasks from cool domains accounted for 85% of the age-related variance on theory-of-mind tasks, and an additional 11.6% of the variance independent of age. In Experiment 3, 3-, 4- and 5-year-olds performed theory-of-mind tasks (false-belief, appearance-reality), and connections and transformations tasks. The connections and transformation items involve similar content to theory-of-mind tasks but have a lower level of complexity, that is they are binary-relational. A predictor task (class inclusion) from a different domain was also included which contained items at two levels of complexity (binary-relational, ternary-relational). Significant complexity effects were found for the theory-of-mind tasks. That is, the binary-relational connections and transformations items were easier and mastered earlier in development than the standard ternary-relational theory-of-mind items. Additionally, the predictor task accounted for more than 80% of age-related variance in theory-of-mind.

Preliminary evidence was also found for a differential age of emergence of ternary-relational processing in hot versus cool domains. As aforementioned, the OFC is thought to develop earlier than DL-PFC (Zelazo & Müller, 2002), hence success on theory-of-mind tasks would be expected to emerge slightly earlier than success on
ternary-relational items of cool tasks. Andrews et al. (2003) found that the majority of children (66% Exp. 1 & 2 combined, \( N = 108 \); and 79% Exp. 2 & 3, \( N = 120 \)) performed consistently, in that they performed significantly above chance level in both hot and cool domains or failed in both domains. The remaining children showed an inconsistent pattern of performance, succeeding in one domain but not the other. Consistent with earlier development of the OFC, these children were more likely to succeed on the theory-of-mind tasks than on transitivity or class inclusion (Andrews et al., 2003). Thus, complexity approaches may provide the means by which the development of hot and cool executive functions can be compared, without confounding the hot-cool distinction with differences in complexity.

*Approaches to the Development of Executive Functions in Childhood*

The recent focus on the development of executive functions in children has raised several practical issues. While there is now considerable evidence that rudimentary executive functions emerge in the preschool years (Hughes, 1998) or even in infancy (Diamond, 2002), there has been a limited range of well-validated assessment methods and tasks designed specifically for children. Instead, many of the currently available measures are downward extensions of adult executive tasks and may therefore not be developmentally appropriate for use with children (Gnys & Willis, 1991). In view of the recent research outlined above (e.g., Andrews et al., 2003; Frye et al., 1995) it may also be beneficial to take into account developmental changes in children’s capacity to deal with task complexity. To date, however, the assessment of executive functioning in children has generally followed the same approaches utilized in the adult literature. While these approaches have typically failed to control for complexity, the overview detailed below serves as a background for the need of a closer examination of the association between developmental changes in executive functioning and task complexity.
Global Assessment of Executive Functions in Children

A number of studies have employed global measures such as the WCST and TOH with school-aged children. Welsh et al. (1991) investigated performance on a battery of executive function tasks, including the WCST and TOH, as a function of age. These investigators hypothesized that given developmentally appropriate behavioural measures, rudimentary forms of prefrontal skills would be exhibited in preschool children. The results indicated that the tasks had different developmental trajectories between 3 and 12 years of age, with adult-level performance achieved at three different ages depending on the nature of the executive function task. Performance on visual search tasks reached adult levels at five years of age, performance on a 3-disc version of the Tower of Hanoi (TOH) test at age six, and on the WCST at age 10. Adult levels of performance on the 4-disc TOH, verbal fluency and a motor sequencing task, however, was not achieved even by 12-year-old children.

With a comparable set of tests, Levin et al. (1991) studied a sample of children aged 7 to 15, divided into three age levels. Major differences were observed between the 7- to 8-year group and the 9- to 10-year group in performance on the WCST and on a go/no-go task. Increments in performance were evident at still higher age levels, between the 9- and 12-year age group and the 13- to 15-year age group, in performance on the California Verbal Learning Test and the TOL test.

In addition, Shum et al. (2000) assessed planning ability in children with traumatic brain injury (TBI) using a 4-disk TOL and the Porteus Maze. Participants were 6- to 10-year-olds ($n = 10$) and 11- to 16-year-olds ($n = 12$) with moderate to severe TBI and unimpaired 6- to 10-year-olds ($n = 15$) and 11- to 16-year-olds ($n = 15$). Five simple TOL problems (requiring two to five moves) and five complex problems (requiring six to nine moves) were administered. Older children performed significantly better than younger children on the 4-disk TOL suggesting that the ability to plan, as
measured by this test, develops with age. Children with TBI performed significantly more poorly than controls on the complex problems but not on the simple problems, while no effect of TBI was found on the Porteus Maze. The authors concluded that children with TBI showed difficulties in planning abilities especially on the complex problems that involved the formation of more subgoals for successful solution.

A number of studies have also utilized these global tasks in order to examine the stage-like emergence of executive abilities. For example, Chelune and Baer (1986) tested 105 children in grades one through six in order to provide developmental norms for the WCST. Children’s performance as measured by the number of categories achieved, perseverative errors, and failures to maintain set, improved most between ages 6 and 7 years, with a somewhat lesser growth spurt occurring between the ages of 8 and 10. The performance of 10-year-old children in terms of categories achieved, was indistinguishable from that of normal adults, whereas perseverative errors did not decline to adult levels until 12 years of age. Anderson, Anderson and Lajoie (1996) administered the TOL to a sample of 376 children aged between 7 and 14 years. The results indicated significant age trends with the older age groups exhibiting shorter solution times and more items completed correctly. The greatest improvements were recorded between the ages of 7 and 9 and again between 11 and 12 years. The number of failed attempts, however, was not related to age.

Both Chelune and Baer (1986) and Anderson et al. (1996) reported that the developmental changes in performance on these global tasks corresponded roughly with both the neuroanatomical changes in the brain and Piaget’s cognitive stages of development as detailed in the previous sections. That is, the development of executive functions is a multiphase process, with different abilities emerging at different times, and maturing in different ways (Stuss, 1992). The question remains, however, whether the emergence of the executive abilities required by these tasks may be explained by
changes in the ability to process task complexity. Criticisms regarding the construct  
validity and reliability of these global tasks as noted in Chapter 1 are also relevant with  
respect to their use in child populations. Performance on these tasks requires a variety of  
other non-executive cognitive processes and as such it is difficult to know exactly why a  
child may experience difficulty at different ages. Thus, several attempts have been made  
to refine these global tasks for use in child samples.

Refinement of Global Tasks

As detailed previously, the DCCS can be considered a simplified version of the  
WCST. Both tasks require children to sort cards according to their shape and colour. In  
the DCCS, however, children are told what the sorting criteria will be, whereas in the  
WCST participants must discover the rule for themselves. The much-replicated finding  
of the DCCS is that 3-year-olds are unable switch to a new sorting rule when instructed  
to do so. Instead, these young children continue to sort cards by the initial sorting  
dimension, even when told the relevant rules on each trial (Zelazo & Frye, 1997).

The IOWA Gambling task has also recently been modified for use with children.  
As described in detail in Chapter 4, Kerr and Zelazo (2004) created a 2-deck version of  
this task that appears suitable for young children. One deck of cards was advantageous  
over trials, while the other deck was disadvantageous. Smarties were substituted for  
money, and the number of gains and losses per card was indicated pictorially by happy  
or sad faces. Kerr and Zelazo (2004) presented 50 trials and found that during the last  
25 trials, 3-year-olds made more disadvantageous choices than would be expected by  
chance, whereas 4-year-olds made fewer. Prior to age 4, children performed like the  
patients with damage to the OFC examined by Bechara et al. (1994).

These refined tasks reduce the issues associated with the task impurity problem,  
however, they are still sufficiently global to capture the breadth of executive functions.  
The DCCS and the Children’s Gambling Task therefore seem to be suitable measures
for use with preschool children. Specifically, they may allow closer examination of the emergence of cool executive functions (as measured by the DCCS) versus hot executive functions (as measured by the Children’s Gambling Task) in children as young as 3-years. Additionally, the complexity of these tasks has been estimated based on both CCC and relational complexity theories. Both of these tasks were therefore utilized in the current research.

**Working Memory Approach to Development of Executive Function in Children**

A number of researchers have attributed the development of executive functioning to changes in the capacity of working memory. As described earlier, Case (1987) proposed that age-related increases in working memory capacity contribute to developmental advances across a broad range of cognitive domains. Consistent with this view, a number of developmental studies have shown that there are regular age-related increases in working memory throughout childhood.

Research conducted by Diamond (1985) and Goldman-Rakic (1987) indicated that the elements of voluntary attention and executive function are formed in the first few years of life. This research utilised the A-not-B task and delayed response tasks. In the A-not-B task, an object is hidden at one location, A, and the infant is allowed to retrieve it. After repeated hiding and retrieval at A, the object is hidden at location B, and the infant is allowed to search for it. The A-not-B error is committed if the infant persists in searching at location A despite having seen the object hidden at location B. Human infants in the 7.5 to 9-month age range and dorsolateral lesioned adult rhesus monkeys avoided the A-not-B error provided that no delay was imposed. However, with delays of 2 to 3 seconds, both groups committed the classic A-not-B error (e.g., Diamond, 1985; Goldman-Rakic, 1987). Espy, Kaufmann, McDiarmid and Glisky (1999) also demonstrated a steady improvement in A-not-B performance in preschool children between 23 and 66 months of age. A constant delay of 10 seconds was used
Development of Executive Functions 45

across trials. Older preschool children retrieved the reward on more trials and made fewer perseverative responses than did younger children, suggesting a continued development beyond late infancy.

Diamond and Doar (1989) investigated the developmental progression on a well-validated measure of spatial working memory, the delayed response task, which involves placing an object under one of two identical covers. After a delay period during which the locations are not visible, the infant is allowed to search for the object. Results indicated that infants were capable of accurately performing a two-choice version of the delayed response task with a 10-second delay as early as 1 year of age. Thus, at 11- to 12-months of age, infants were already able to carry out an intention on the basis of stored information. However, the sensitivity of such forced two-choice delayed response tasks to subtle differences in working memory has been questioned because of the ease with which such tasks are performed by most individuals over the age of one year. More difficult versions of the spatial delayed response paradigm have since been utilized. Zald and Iacono (1998), for example, assessed the spatial working memory abilities of 14- to 20-year-olds on a multi target spatial delayed response task. Performance improved significantly from 14 to 20 years of age, suggesting continued development during the second decade of life. However, this may be attributable to increasing ability to process the level of complexity required by the more difficult task.

Luciana and Nelson (1998) administered measures of spatial memory span, spatial working memory, the TOL, visual pattern and spatial recognition tasks, and a set shifting task to a sample of 4- to 8-year-old children, as well as a small group of young adults. The results indicated an age-related progression with 4-year-olds performing worse than 5- to 7-year olds on all tasks. Even 8-year-old children, however, had not reached adult levels of performance on the complex problems such as the TOL and spatial working memory tasks. The authors concluded that the development may occur
dimensionally, commencing with basic perceptual and sensorimotor functions and culminating with complex processing required by working memory tasks.

Case, Kurland and Goldberg (1982) developed a counting span task that is analogous to the reading span task devised by Daneman and Carpenter (1980). This task consists of a processing component (counting the number of coloured dots on a computer screen) and a storage component (maintaining the products of these counting operations). Case et al. measured counting span among 6- to 12-year-old children and, also separately measured children’s counting speed in order to obtain a measure of processing efficiency. Counting span increased during development, and the amount of growth was predicted by age-related increases in the speed of processing. Similar to Daneman and Carpenter (1980), memory span capacity was presumed to reflect the capacity to share resources between the dual functions of processing and storage. However, as children develop, processing is thought to become faster and more automatic, thus freeing resources for storage and hence increasing memory span.

Gathercole, Pickering, Ambridge and Wearing (2004) examined the performance of over 700 children between the ages of 4 and 15 years on measures associated with the three major components of Baddeley’s (1974) working memory model. The results indicated that the capacity of the phonological loop (as measured by digit recall, word recall, and nonword recall), visuospatial sketchpad (block recall, memory for route, Visual Patterns test) and central executive (measured by backward digit recall, and two complex span tasks, listening recall and counting recall) each increased linearly from age 4 to early adolescence. Additionally, a confirmatory factor analysis that consisted of three factors that corresponded to Baddeley’s working memory model provided a good fit of the data from 6 years of age onwards. Thus, the structural organisation of working memory not only appeared to remain constant over the childhood years, but also reflected the tripartite structure of adult working memory.
Inhibitory Approach to Development of Executive Function in Children

Early efforts to explain the developmental progression in executive skills focused on the role of inhibition (e.g., Luria, 1959). Luria reported that damage to the prefrontal cortex produced disinhibition in attention, memory, reasoning and problem solving. The maturation of the prefrontal cortex was consequently hypothesized to lead to increased inhibitory control, with age-related changes in executive functions being attributed to the growth of an inhibitory mechanism (Dempster, 1992). Individuals with executive function deficits were thought to try to suppress interfering response tendencies but were unable to do so because of an immature or inefficient mechanism (Dempster, 1992). Research investigating the development of inhibition in children has largely focused on two key aspects, response inhibition and switching processes that involve conflict.

Bell and Livesey (1985) investigated the development of response inhibition in children aged three to six years using the go/no-go task. The study found that although young children understood the significance of discriminative cues in the task, they were unable to withhold motor responding. The authors concluded that children younger than six years of age were able to verbalise when a response should be inhibited but they could not necessarily make the related motor response.

Diamond, Kirkham and Amso (2002) investigated conflict performance on a simplified version of the Stroop task, the day-night task. In the original version of this task, children were presented with either a day picture of the sun, or a night picture of the moon (and some stars) and were instructed to say ‘night’ to the day picture and ‘day’ to the night picture. Diamond et al. (2002) added a second condition in which children were instructed to say ‘dog’ to the day picture and ‘pig’ to the night picture. The authors found that 4-year-old children responded more accurately in dog-pig condition than the standard day-night condition. The problem for the children appeared
Development of Executive Functions

to lie in inhibiting a word that is semantically related to the word they are trying to say. Inhibitory demands were reduced in the dog-pig condition because the correct response was semantically unrelated to the to-be-inhibited response, thereby eliminating conflict.

Espy (1997) investigated age-related changes in both response inhibition and switching processes. The performance of 70 preschool children (32-68 months of age) was examined on a new task, the Shape School. The task was presented in a colourful storybook format that depicted characters that differed in shape (circle or square) and colour (red, blue or yellow). The task included three conditions: control, inhibition and switch trials. In the control condition, children were required to name the characters according to their colour. In the inhibition condition, children were required to name only the colours of the characters that had a smiling (not neutral) face. In the switch condition, another set of characters wearing hats was added to the story. This phase involved conflict as children were required to name the colours of the characters without hats, but to name the shapes of the characters that were wearing hats. The results indicated that older age groups performed more efficiently than younger groups. More specifically, inhibition efficiency improved significantly between three to four years of age, whereas switching skills showed developmental improvement from four to five years of age. Thus, response inhibition skills may develop earlier than the switching abilities required for dealing with conflicting responses.

Carlson and Moses (2001) examined the performance of 107, 3- and 4-year-old children on theory-of-mind tasks and two types of inhibitory tasks: delay tasks and conflict tasks. The four delay tasks involved measuring a child’s ability to delay or suppress an impulsive response to particular stimuli. This was similar to the inhibition condition of the Shape School. The six conflict tasks, including the DCCS, required children to switch flexibly between different responses, depending on task rules or instructions, in a similar manner to the switch condition of the Shape School. Consistent
with the results of Espy (1997), 4-year-old children generally outperformed 3-year-old children on both types of inhibitory tasks. TOM performance was significantly correlated with both conflict and delay tasks, however, the conflict tasks related more strongly to TOM than did the delay tasks.

Interpretations of the results of both Espy (1997) and Carlson and Moses (2001) have contrasted the performance of children on the inhibition versus switching trials. These studies have made a qualitative distinction between these two processes, and the development of inhibition skills were suggested to occur earlier than the emergence of switching abilities. However, it remains to be seen whether the items assessing these two processes are comparable in terms of the complexity of the processes required for success. For example, the level of difficulty required by the inhibition trials may be lower than the level required by switching trials, thereby allowing an earlier age of attainment. The benefit of complexity analyses of the processes underlying the performance of executive function tasks is that they allow for the differential comparison of the emergence of each process without confounding complexity.

As with the adult literature, however, studies that emphasise just one aspect of functioning may be too simple to capture the full range of skills covered by the executive function domain. Further, single construct approaches such as inhibition do not yield testable predictions regarding which situations will be most likely to pose difficulty for children at varying ages (Zelazo & Müller, 2002). Models that incorporate several aspects of the executive function domain may be more appropriate in the assessment of executive functions.

Multiple Component Approaches

Combined inhibition/working memory explanations. A third approach to the development of executive processes has been to suggest that intact executive functions involve both working memory and inhibition. As outlined earlier, Roberts and
Pennington (1996) proposed that inhibition and working memory are most often interactive, with working memory load influencing the extent of inhibitory responses (Roberts & Pennington, 1996). However, if the prepotent response is particularly strong or the working memory demand high, cognitive processes break down and produce strong perseverative responses. Although this proposal is more complex than accounts that focus on inhibition or working memory alone, Zelazo and Müller (2002) suggest that it is still a fairly simple model and the efficiency of this model in accounting for the full range of executive functioning constructs remains unclear. Beveridge, Jarrold and Pettit (2002) provided a test of this model by administering three executive functions tasks, each of which systematically combined two levels of memory load and two levels of inhibitory demands, to 6- and 8-year-old children. While significant effects of memory and inhibition were independently observed, there was little evidence to suggest that these two variables interacted in any way. The authors concluded that memory load and inhibitory demands did not draw upon a shared pool of executive function resources, therefore failing to support the interactive model proposed by Roberts and Pennington (1996).

Multiple component by statistical analysis. A number of developmental studies have found evidence for the non-unitary nature of executive functions. As described above, Welsh et al. (1991) examined the performance of 3- to 12-year-olds on a battery of executive function tasks including the WCST, TOH, visual search, verbal fluency and a motor sequencing task. A principal component analysis indicated that the individual measures clustered into three different factors: speeded responding, set maintenance, and planning. Similarly, Levin et al. (1991) examined the performance of children aged 7 to 15 on the WCST, TOL, California Verbal Learning test and a go/no-go task. Results also revealed a three-factor solution that included concept formation, freedom from perseveration, and planning. Zelazo and Müller (2002) cautioned that the
interpretations of such results are potentially ambiguous. The factors that emerge depend to a large extent on which tasks are included in the test battery, scoring processes and dependent measures used, and which variables are entered into the factor analyses. The labels assigned to the various factors may also not align with the cognitive processes underlying performance on the various tasks.

Lehto, Juujärvi, Kooistra, and Pulkkinen (2003) used confirmatory factor analysis techniques to examine the dimensions of executive functioning in 8- to 13-year-old children. One hundred and eight children were administered a battery of eight executive function tasks. Similar to the results obtained by Miyake et al. (2000) which were described in the previous chapter, factor analyses yielded three interrelated factors. The factors were labelled Working Memory (Auditory Attention and Response Sets-Part B, Mazes, Spatial Working Memory Task, Spatial Span), Inhibition (Matching Familiar Figures, TOL) and Shifting (Trail-Making Test-Part B, Word Fluency). While performance across the eight executive tests was best accounted for by these three separate factors, significant correlations among the executive function factors were also found.

Multiple components in standardised test batteries. A number of standardized test batteries have been developed to assess the progression of children across a range of these executive abilities. For example, the NEPSY (A Developmental Neuropsychological Assessment; Korkman, Kemp & Kirk, 1988) is used to assess the neuropsychological development of children between the ages of 3 and 12 years. The NEPSY assesses performance across five core domains: attention/executive functions, language, visual-spatial processing, sensorimotor, and memory and learning. The attention and executive function domain includes six subtests: visual search, statue, tower, auditory attention, design fluency and knock-and-tap. Individual performance is summarized in test profiles that demonstrate children’s strengths and weaknesses. The
Development of Executive Functions

test battery was normed on a sample that included 100 children for each age year between 3 and 12 (N = 1000) with average reliabilities (including split-half and test-retest) across the domains ranging from 0.70 to 0.91.

The appeal of such standardized test batteries lies in the normative data that allows for comparison of individual results to age-normed levels of performance. While batteries such as the NEPSY therefore have utility as criterion measures of executive functioning, they do not allow for a distinction to be made between the development of hot versus cool tasks. The NEPSY also makes comparisons across different age groups difficult, as it utilizes different subtests and different items between the different ages. Additionally, the complexity of the underlying cognitive processes required for success on each individual task is not taken into account when test profiles are interpreted.

Problem solving approach to executive functioning. As aforementioned, micro-level approaches that highlight a particular subfunction may not adequately capture the higher-order processes that constitute executive control. Zelazo, Carter, Reznick and Frye (1997) therefore proposed a problem-solving framework that treats executive function as a macro construct. This descriptive framework delineates four temporally and functionally distinct phases of problem solving. That is, executive function involves (i) representing a problem flexibly, (ii) planning organized sequences of thought or action, (iii) executing those sequences, and (iv) evaluating the results of one’s rule use. For example, successful performance of the WCST requires constructing a representation of the problem space, including identifying the relevant dimensions, planning an appropriate course of action, such as sorting by colour, carrying out the prescribed action, and evaluating the outcome of one’s behaviour, including using feedback to detect errors. Dramatic changes in all four aspects of executive function occur between two and five years of age as children gradually come to control their thought and behaviour during problem solving (Zelazo et al., 1997). Disinhibition, as
evidenced by perseveration, can also occur in any of these four phases of problem solving. While the framework does not explain executive functions, the authors suggest that it allows for a closer examination of exactly where in the process of problem solving performance breaks down (Zelazo et al., 1991).

Relational Complexity Approach to Executive Function

According to relational complexity theory, many higher cognitive processes (including executive functions) involve the processing of relations. For example, planning as measured by the TOH and other problem solving tasks involve the relation between current states and goal states (Halford et al., 1998). Inhibitory control, which is frequently described as a core executive function (e.g., Carlson & Moses, 2001) is suggested to involve the processing of relations because inhibitory control tasks often require switching between alternative stimulus-response mappings (Andrews et al., 2003). The WCST is suggested to involve the processing of relations, as planning to sort cards by the alternate rule adds a further element of complexity to the task. The representation of relations is at the core of analogies (Halford et al., 1998), and, therefore, tasks that require analogical reasoning and analogical transfer to novel situations also involve processing relations.

Relational processing can be distinguished from processing based on associations. For example, unlike processing based on association, relational processing can involve a type of flexibility known as omnidirectional access, which means that given all but one of the components of a relational instance, the remaining component can be retrieved (Halford et al., 1998). Relational processing is also symbolic, with links between the arguments of a relation explicitly symbolized (e.g., one link between “whale” and “dolphin” is explicitly symbolized by “bigger-than”). A further property of relations is the ability to transfer to isomorphs, which enables an individual to deal with novel situations. The processing of relations can also be described as effortful, because
processing loads increase with task complexity. This is represented in the relational complexity metric described previously. Relational processing can therefore be described as flexible, symbolic, effortful and as involving planning and coping with novel situations. Thus, there is considerable overlap between these properties of relations and the characteristics of executive functions outlined in Chapter 1.

As aforementioned, the relational complexity metric has been shown to provide an effective measure of cognitive development in cool content domains including transitivity, hierarchical classification, class inclusion, cardinality, hypothesis testing and comprehension of relative-clause sentences (e.g., Andrews & Halford, 2002), balance scale (Halford, Andrews, Dalton, Boag, & Zielinski, 2002) and property inference based on classification hierarchies (Halford, Andrews & Jensen, 2002). With the exception of theory of mind (e.g., Andrews et al., 2003), little research has been conducted into the development of hot executive functions, and fewer studies have compared the emergence of hot versus cool executive function skills. Investigations into the development of executive functions, including distinctions between hot and cool domains, will continue to benefit from taking into account the complexity of the cognitive processes that the tasks require.

Conclusions

It can be concluded that future research into the development of executive functions needs to take several factors into consideration. Firstly, the use of refined versions of global tasks, such as the DCCS and the Children’s Gambling Task, may enable researchers to counter criticisms related to construct validity due to task impurity by specifically identifying the cognitive processes involved in task performance, while also capturing the breadth of the executive function construct. Secondly, adequate assessment of executive functions needs to sample both the hot domain associated with the OFC and the cool domain associated with the DL-PFC. Thirdly, many important
changes in the development of executive functions occur between the ages of three and six years. Assessment of this age range needs to ensure that developmentally appropriate measures are utilized. According to relational complexity theory, this would involve tasks that are presented at binary- and ternary-relational levels. Fourthly, future investigations regarding differential impairments in executive function, either between hot versus cool domains or across clinical groups, need to take complexity into account.
CHAPTER 3. OVERVIEW OF THE EMPIRICAL RESEARCH

The current research aimed to further investigate the development of executive functions in 3- to 6-year-old children. Two broad domains of executive functions were targeted: cool executive functions associated with the DL-PFC and hot executive functions associated with the OFC. The study intended to use developmentally appropriate tasks that were presented at a complexity level that was appropriate for the targeted age groups, that is binary- and ternary-relational items. Further, the study aimed to ensure that any hot-cool comparisons were not confounded with task complexity.

The current study extends previous research by providing a more comprehensive test of the applicability of the relational complexity metric to hot executive functions by examining three new domains: complex decision making, conditional discrimination and future-oriented decision making. Complex decision making, to be described in Chapter 4, was assessed using the Children’s Gambling Task. Future-oriented decision making, to be detailed in Chapter 5, places participants in conflict situations, for example where they must choose between receiving a less desirable reward now and receiving a more desirable reward later. Chapter 6 will describe the conditional discrimination/reversal-learning domain, which required children to either make a choice between two geometrical stimuli (reversal learning) or the choice between two stimuli was made conditional on a further cue, such as the background on which they were displayed (conditional discrimination).

Previous research points to the involvement of the OFC in the performance of complex decision making (e.g., Bechara et al., 1994), future-oriented decision making (e.g., Zelazo & Müller, 2002) and conditional discrimination/reversal learning (e.g., Overman, 2004). For each of these hot executive functions an existing task will be described. Then, a relational complexity analysis of that task will be presented. Based
on these complexity analyses, binary-relational and ternary-relation items of each of these tasks were developed or existing tasks were selected and/or modified. The binary-relational items were closely matched to the ternary-relational items in terms of stimuli and procedure, however, they were lower in complexity.

The three measures of hot executive functioning were first included in a pilot study. Items at both level of complexity for each task were administered to 3-, 4- and 5-year-old children. The results of the pilot study will be reported within each relevant task chapter (i.e., chapters 4, 5 & 6).

Table 3.1

Selection of Tasks in Pilot and Main Study from Hot and Cool Domains of Executive Function

<table>
<thead>
<tr>
<th>Study</th>
<th>Hot Executive Tasks</th>
<th>Cool Executive Tasks</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot</td>
<td>Children’s Gambling Task</td>
<td></td>
<td>CD/RL</td>
</tr>
<tr>
<td></td>
<td>Future-oriented decision making</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main</td>
<td>Children’s Gambling Task</td>
<td>Transitivity</td>
<td>NEPSY</td>
</tr>
<tr>
<td></td>
<td>Future-oriented decision making</td>
<td>Class Inclusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD/RL</td>
<td>DCCS task</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theory of Mind</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be seen in Table 3.1, the three new measures of hot executive functioning were then included in the main study that was administered to a sample of normally developing 3-, 4-, 5- and 6-year-old children. Existing binary- and ternary-relational theory-of-mind items were also administered (taken from Andrews et al., 2003), which
required children to recognize mental states, such as beliefs, desires, intentions and emotions, in themselves and others. Additionally three cool measures were selected, each having previously been analysed in terms of their relational complexity: (a) transitivity, which requires reasoning about ordinal spatial relations, (b) class inclusion, which involves reasoning about inclusion hierarchies, and (c) a representational rule use task, the DCCS task. All children in each age group completed all tasks. The inclusion of measures of both hot and cool executive functions, each with complexity manipulated, was intended to allow for the examination of a possible differential age of emergence of executive abilities associated with the DL-PFC versus the OFC.

A selection of subtests from the Attention and Executive Function sub-domain of the NEPSY was also administered in the main study. The NEPSY is a standardized psychometric test battery (Korkman, 1999), which was utilized as a reference test to compare levels of functioning in the current sample with previous research.

General Method

Participants

Pilot study. Seventy-two children from three day-care centers participated. Three age groups were recruited with equal numbers of males and females: 24, 3-year-olds ($M = 44.58$ months, $SD = 2.32$, range = 39-47), 24, 4-year-olds ($M = 52.42$, $SD = 2.52$, range = 48-58), and 24, 5-year-olds ($M = 64.83$, $SD = 3.78$, range = 60-71).

Main study. One hundred and twenty children from two day-care centers and one primary school participated. Children from four age groups were recruited with equal numbers of males and females: 30, 3-year-olds ($M = 42.60$, $SD = 3.56$, range = 36-47), 30, 4-year-olds ($M = 52.87$, $SD = 2.50$, range = 49-57), 30, 5-year-olds ($M = 65.90$, $SD = 3.57$, range = 60-71), and 30, 6-year-olds ($M = 77.10$, $SD = 2.76$, range = 72-83).
The participants were recruited through letters distributed to parents by the day-care centers and school staff (see Appendix A). Written consent was obtained from parents prior to testing, and recruitment and task procedures were conducted in accordance with specific guidelines established during ethics approval applications from the Griffith University Human Research Ethics Committee (APY/78/04/HREC).

An additional three children completed some tasks but were not included in the main study. In each of these cases, the child’s parents or a staff member informed the experimenter of a pre-diagnosed developmental disorder. These children were a 3-year-old female with an unspecified developmental delay, a 4-year-old male with cerebral palsy, and a 5-year-old male with attention deficit disorder. In each case, parental consent had been obtained, and the experimenter did not want the children to feel excluded in any way. While the complete test battery was not administered to the children, they did complete a number of tasks in order to make sure they had received as many rewards (stickers and chocolates) as the other participating children.

Representativeness of the Sample

In the main study, each age group was also administered the relevant age-appropriate sub-tests for the NEPSY executive function and attention domain. The statue and visual attention tasks were administered to 3- and 4-year-olds while the 5- and 6-year-olds completed the visual attention, tower and auditory attention tasks. The visual attention test items differed for 3- and 4-year-olds versus 5-and 6-year-olds. The mean performances of each age group on the appropriate subtests are presented in Table 3.2.
Table 3.2

*Mean (SD) Performance on NEPSY Subtests Across Age Groups (N = 120)*

<table>
<thead>
<tr>
<th>Sub-test</th>
<th>Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Statue</td>
<td>20.67</td>
</tr>
<tr>
<td></td>
<td>(4.17)</td>
</tr>
<tr>
<td>Visual Attention</td>
<td>13.10</td>
</tr>
<tr>
<td></td>
<td>(5.73)</td>
</tr>
<tr>
<td>Tower</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Auditory Attention</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Children’s scaled scores for the subtests were combined to produce an executive function and attention domain score. The mean domain scores of each age group are presented in Table 3.3. This table includes the mean percentile equivalents for 3- to 6-year-olds. The mean of all age groups was at or above the 70th percentile and ranged between the 25th and 99th percentile. The performance on the NEPSY indicates that the sample of children in the current study was representative of the normal population.
Table 3.3

*Mean (SD) Executive Domain Scores and Percentile Rankings Across Age Groups in the Main Study*

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Domain Score (SD)</th>
<th>Mean Percentile Rankings</th>
<th>Percentile Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>23.67 (2.59)</td>
<td>72.86 (15.69)</td>
<td>42.00 – 99.00</td>
</tr>
<tr>
<td>4</td>
<td>22.90 (2.52)</td>
<td>70.10 (17.44)</td>
<td>27.00 – 90.00</td>
</tr>
<tr>
<td>5</td>
<td>34.63 (3.99)</td>
<td>75.00 (20.03)</td>
<td>25.00 – 99.00</td>
</tr>
<tr>
<td>6</td>
<td>36.83 (3.87)</td>
<td>83.87 (15.62)</td>
<td>45.00 – 99.00</td>
</tr>
</tbody>
</table>

*Procedure*

Children were tested individually by a female experimenter in a quiet area of the day-care center or school, separated from the other children. As outlined in Table 3.1, the pilot study included the administration of three hot executive function tasks. The main study included seven tests, four in the hot executive function domain and three in the cool executive function domain. Tasks were administered as described in the subsequent chapters to all children. The presentation order of the seven tasks, and the complexity levels within each task were counterbalanced. Total testing time ranged from approximately 1 hour 30 minutes for the oldest child, to approximately 2 hours 45 minutes for the youngest children. Testing was completed during two to four sessions depending on the child’s age, motivation, attention level and school schedule. At the beginning of each session, the children were told, “*We’re going to play some games where I will show you some things, and tell you some stories, and then ask you some questions about them. Is that all right with you?”* The child’s assent was obtained prior
to the commencement of testing. Responses to all tasks were recorded on response sheets.

Treatment of Results

The results for each task in the relational complexity battery are presented following the introduction and method for each task. To the extent that the relational complexity metric applies to both hot and cool domains, items of higher relational complexity (ternary-relational items) should be more difficult than items of lower relational complexity (binary-relational items). Performance across all tasks should also improve with age. In accordance with the relational complexity approach, however, age effects should be strongest on the ternary-relational items. Complexity and age effects and interactions of complexity with age were assessed using a 2(complexity) × 4(age-group) × 2(gender) factorial analysis of variance (ANOVA), except when assumptions of ANOVA were violated. When this occurred nonparametric tests were substituted. An alpha level of .05 was used for all analyses. The assumptions of the individual analyses were met unless otherwise specified.

Single sample $t$ tests were also used to compare the mean level of performance for each age group to chance level, that is the level of performance that would result if the children guessed. Individual pass-fail scores were also computed. The percentage of children in each age group whose scores exceeded the chance baseline for the binary- and ternary-relational items within each domain was also calculated. It was expected that less than 50% of the 3-year-olds would pass the ternary-relational items, whereas more than 50% of the 5-year-olds and 6-year-olds would pass. The percentage for 4–year-olds would fall in between, because this is the transitional period. These results are reported in Chapters 4 through to 7 for the hot executive function domains and Chapter 8 for the cool executive function tasks.
Chapter 9 examined the relationship between performance on the tasks associated with the hot and cool executive function domains. Functions associated with the OFC were expected to emerge earlier than those associated with the DL-PFC. Analyses from the previous chapters were combined to examine the patterns of age-related improvements and pass-fail percentages on the ternary-relational items from each domain. Correlational analyses were performed to investigate the within and across domain associations. Factor scores for hot executive functions and cool executive functions were computed, and the effects of age and domain were examined.
CHAPTER 4. COMPLEX DECISION-MAKING

One task that falls under the definition of hot executive functions, and that is sensitive to deficits of the OFC, is the Iowa Gambling Task developed by Bechara et al. (1994). This decision-making task requires participants to make a series of card selections that result in winning and losing money. Participants were given an initial stake of $2000 (play money) and instructed to win as much money as possible by choosing cards from any of the four decks. The four card decks are characterized by different reward-punishment profiles, such that decks A and B offer high rewards but higher penalties, resulting in overall loss, whereas decks C and D offer smaller rewards but minimal penalties, resulting in overall profit. Unimpaired adults are quick to identify the advantageous decks and to select cards from these decks, while avoiding the disadvantageous decks. The proportion of cards chosen from the advantageous decks increases across trials. In contrast, Bechara et al. (1994) found that patients with orbitofrontal lesions are guided by immediate prospects rather than long-term consequences. Even when they can describe the contingencies governing the task, these patients cannot use this knowledge to guide their behaviour. They continue to select cards from the disadvantageous decks.

The Somatic Marker hypothesis was proposed in order to provide a neural explanation of the difficulties encountered on gambling tasks by patients with OFC lesions. This hypothesis postulates that decision-making is a process guided by emotions, and that there may be an association between abnormalities in emotion in these patients and their impairments in decision-making tasks (Bechara, 2004). More specifically, it is argued that the OFC (or ventromedial PFC) is part of a network involved in associating somatic (bodily) states with various situations. Repeated encounters with the same situations leads to a reactivation of the previous somatic state
(Damasio, 1994). However, the development of this anticipatory response may be lacking in the patients with OFC lesions.

In order to test this hypothesis, Bechara et al. (1994) examined the performance of both unimpaired controls and ventromedial patients on the gambling task while recording their skin conductance responses (SCRs). The results indicated that both normal controls and OFC patients generated SCRs after they had picked a card and been told that they won or lost money. As normal controls became familiar with the task, they began to generate SCRs prior to the selection of any cards, that is, they developed anticipatory SCRs which were more pronounced before selecting from the riskier decks. In contrast, the lesioned patients entirely failed to generate any SCRs before picking a card. Therefore successful completion of decision-making task may require guidance by emotional signals that are generated in anticipation of future events.

Rogers et al. (1999) utilized a computerized decision-making task, in which participants first made a fairly simple probabilistic decision, and then gambled points on their confidence in this decision. The task minimized working memory and learning processes by visually presenting all necessary information on the screen. Patients with OFC lesions took considerably longer to make their decisions, and were more likely to make poor choices compared to age- and IQ-matched controls. In contrast, patients with dorsolateral or dorsomedial lesions were unimpaired.

A number of functional neuroimaging studies have also been conducted which have indicated activation in the OFC in processing rewards and punishments (e.g., Elliot, Dolan & Frith, 2000; Elliot, Frith & Dolan, 1997). For example, O’Doherty, Kringelbach, Rolls, Hornak and Andrews (2001) used fMRI to investigate performance on a probabilistic visual association task in which participants could win and lose money. The results indicated that the OFC was not only activated by monetary rewards and punishment, but that the magnitude of the activation was correlated with the
magnitude of the rewards and punishments received. The authors concluded that the OFC may be involved in emotional processing by representing the magnitudes of abstract rewards and punishments, such as receiving or losing money.

Gehring and Willoughby (2002) recorded ERPs from human participants as they performed a monetary gambling task. Participants’ choices on this task were followed by outcome events that signified not only the amount of money won (or lost), but also the loss (or gain) that would have resulted from the alternative choice. The results indicated ERPs in the medial frontal region, which was greater in amplitude when a participant’s choice between two alternatives resulted in a loss than when it resulted in a gain. Combined with the previous imaging and lesion studies, the results suggest a significant role of the OFC in processing rewards and punishments, and in particular in monetary gains and losses.

Gambling Tasks in Child Populations

The IOWA Gambling task has recently been used in developmental samples. For example, Crone and van der Molen (2004) examined the performance of children, adolescents and young adults from 6 to 25 years of age on computerized versions of the task. Prior to age 12, the results indicated that children opted for choices that yielded high immediate gains in spite of their higher future losses. The results of these younger children modeled the behavioural patterns displayed by patients with damage to the OFC described above.

Garon and Moore (2004) modified the IOWA Gambling task for use with even younger children. They substituted Smarties for money, and reduced the number of trials from 100 to 40, but they retained four decks of cards. They examined the performance of 3-, 4-, and 6-year-old children. The children were also given an awareness test to assess their understanding of what was happening in the game.
Awareness improved with age. However, there was no significant effect of age on selections from advantageous decks across the blocks of trials, and even 6-year-old children appeared to perform at chance levels only.

Kerr and Zelazo (2004) created a 2-deck version of the task that seems more suitable for young children. When turned, the cards displayed a number of happy and sad faces. Happy faces indicated the number of candies won, while sad faces indicated the number of candies lost. One deck of cards was advantageous over trials, and the other deck was disadvantageous. Consistently choosing from the advantageous deck resulted in a net gain of rewards, whereas consistently choosing from the disadvantageous deck resulted in a net loss. Within each deck, the number of gains was constant across cards, but the number of losses was variable. Cards in the advantageous deck always provided a gain of one reward (i.e., one happy face) together with either zero or one losses. Selections from this deck resulted in a net gain of 5 candies per block of 10 trials. Cards in the disadvantageous deck always provided a gain of 2 rewards together with losses of 0, 4, 5, or 6 candies. Selections from this deck resulted in a net loss of 5 candies per 10 trials. Kerr and Zelazo (2004) presented 50 trials and found that during the last 25 trials, 3-year-olds made more disadvantageous choices than 4-year-olds, and 3-year-olds made more such choices than would be expected by chance, whereas 4-year-olds made fewer. Prior to age four, children performed like patients with damage to the OFC as found by Bechara et al. (1994).

**Complexity Analysis of the Children’s Gambling Task**

Kerr and Zelazo (2004) interpreted their results in terms of the CCC theory. As noted in Chapter 2, this approach proposes age-related changes in the complexity of the rule systems that people are able to use permit children to exercise increasing control over their thoughts and actions. With respect to the Children’s Gambling Task, gains and losses are put into conflict. Consideration of only the gains will result in children
selecting from the disadvantageous deck, however, consideration of both gains and losses produces selections from the advantageous deck. Integrating gains and losses requires the consideration of two dimensions in contradistinction (Kerr & Zelazo, 2004). Thus 3-year-olds are able to learn the initial discrimination between the decks, for example “the striped deck leads to high gains, while the dotted deck leads to low gains”. However, these younger children have difficulty coordinating this initial discrimination with emerging evidence regarding losses (e.g., striped deck has high losses, dotted deck has low losses). In contrast, older children are better able to reflect on these conflicting discriminations and formulate a higher-order rule that allows them to appreciate net gains (Kerr & Zelazo, 2004).

Kerr and Zelazo’s (2004) results can also be accounted for by Relational Complexity Theory. According to this approach, the Children’s Gambling task is ternary-relational. It requires computation of the difference between the decks in wins and losses, and the integration of these differences. This can be thought of as the integration of two binary relations, involving three variables (Deck, magnitude of gain, magnitude of loss). According to relational complexity theory, 3-year-olds will be able to process the component binary relations but they will be unable to integrate these binary relations into a ternary relation. It is not until approximately five years of age that children are able to process the ternary relations require for successful completion of the task.

For the current study, less complex binary-relational versions of the Children’s Gambling Task were created. The binary-relational versions were closely matched to the ternary-relational items in terms of stimuli and procedure, however, complexity was lower. In the binary-relational versions, the decks differed in either gains only or losses only, with the other variable held constant. Specifically, in the binary-gain version, the two decks varied in terms of the magnitude of the gain only, while loss was held
constant across the decks. In the binary-loss version, the two decks varied in terms of
the magnitude of the loss, while the magnitude of the gain was held constant across the
decks. In both the binary-gain and binary-loss versions, the child’s task was to learn to
choose the advantageous deck.

The original ternary-relation version and the two newly created binary-relational
versions were administered to children in both the pilot and main studies. Children’s
performance on the ternary-relational version was anticipated to be consistent with Kerr
and Zelazo (2004) findings and with the predictions of relational complexity theory.
That is, 5-year-olds should be able to integrate three variables (Deck, magnitude of the
gain, magnitude of the loss) and consequently make more advantageous choices across
trial blocks than would be expected by chance. On the other hand, 3-year-olds should be
able to process the component binary relations involved in this task, but they should be
unable to integrate these binary relations into the ternary relation required for successful
completion. Thus, the 3-year-old children would make fewer advantageous choices
across trial blocks than would be expected by chance. The performance of 4-year-olds
was anticipated to fall in between that of the 3-year-olds and that of the 5- and 6-year-
olds.

The newly created binary-relational versions of the Gambling Task, however,
only require the consideration of just two variables (Deck, magnitude of gain) or (Deck,
magnitude of loss). As such, all four age groups should succeed and make more
advantageous choices across trial blocks than would be expected by chance. Success on
binary–relational versions accompanied by failure on the ternary-relational version
would indicate that 3-year-olds’ difficulty was due to the complexity of the task rather
than to the aspects of the materials or task procedures.
Method

Materials.

There were three sets of laminated cards, one set for the ternary-relational version and two sets for the binary-relational versions. Each set contained two decks of 50 cards. For ease of identification, all cards in a single deck had the same pattern on the reverse side. The pattern differed across decks. For example, the reverse sides of cards in ternary-relational deck A were covered with black dots, whereas the reverse sides of cards in deck B were covered with black vertical stripes.

The front-sides of all cards were divided into a white upper half and a black lower half. Black happy faces appeared on the upper half of the card, and white sad faces appeared on the lower half (e.g., Figure 4.1). The number of happy faces on each card indicated the number of rewards gained, and the number of sad faces indicated the number of rewards lost. The bottom half of each card was covered with a yellow Post-It note paper, which the experimenter raised in order to reveal the sad faces beneath.

![Sample card face from the ternary-relational disadvantageous deck of the Children’s Gambling Task (gain = 2, loss = -4).](image)
For the ternary-relational version, the number of gains was held constant across cards within each deck, but the number of losses varied. Cards in the advantageous deck always provided a gain of one reward (i.e., one happy face) together with zero or one loss (with a net gain of 5 rewards per block of 10 trials). Cards in the disadvantageous deck always provided a gain of two rewards together with losses of 0, 4, 5, or 6 candies/stickers (with a net loss of 5 candies per 10 trials). The order of the cards in each deck and the gain/loss contingencies were identical to Kerr and Zelazo (2004) and proportional to those used by Bechara et al. (1994).

In the binary-gain version, the two decks differed in terms of the magnitude of the gain only, while the magnitude of the loss was held constant across the decks. Cards in Deck A provided a gain of one reward on every trial, with a loss of one reward on 50% of trials. Cards in Deck B provided a gain of two rewards on every trial, with a loss of one reward on 50% of the trials. In the binary-loss version, the two decks differed in terms of the magnitude of the loss only, while the magnitude of the gain was held constant across the decks. Cards in Deck A always provided a gain of one reward, with a loss of one reward on 50% of trials. Cards in Deck B also provided a gain of one reward on every trial, however, the magnitude of the loss was greater. Children lost five rewards on 50% of the trials.

Rewards consisted of mini M&M chocolates (or stickers, depending on the expressed wishes of the child’s parent). When children gained rewards, they were placed in a transparent container positioned in front of the child. When children lost rewards, they were removed from the child’s container and returned to an opaque container in front of the experimenter. The response sheets used for recording each child’s selections are presented in Appendix B.
Development of Executive Functions

Procedure

_Pilot study._ Task procedures were based on Kerr and Zelazo (2004), and identical procedures were used for binary- and ternary-relational versions. The three versions were presented in a counterbalanced order. Each version consisted of six demonstration trials and 50 test trials. Each child was initially given one reward to motivate them to play. They were then introduced to the task and given ten chocolates / stickers with which to play the game. The experimenter demonstrated the game by selecting three cards consecutively from each deck. Explanations of rewards and losses were provided. For example “*Look, there are two happy faces – that means that you win two M&Ms*.“ Two candies were placed on the happy faces and then deposited in the child’s container. The experimenter then checked for losses, “*Okay, now we have to see if there are any sad faces. Oh look, there are four sad faces – that means you lose four M&Ms*.“ The experimenter then removed four M&Ms. At this point, the experimenter said, “*We don’t like sad faces do we, because we lose M&Ms. But we like happy faces, right, because we win M&Ms!*”

After the demonstration trials, the experimenter said “*Okay, now we’re ready to start playing the game. You get to choose whichever card you want to play with every time. You can play from the dots or the stripes or from both. You get to choose one card every time and you can pick as many cards as you want until I say STOP and then the game will be done. So remember, you want to make sure that you win as many M&Ms as possible! Let’s see if we can fill this container right up to the top with chocolates (stickers)! Whatever you have in the container at the end of the game you can eat. Okay? Which card do you want to pick first?*”

Test trials were administered exactly like the demonstration trials. The children were not allowed to eat any of the candies (or play with the stickers) they won until after the last test trial (trial 50). Children were not told how many trials there would be.
Five blocks of ten trials were administered for each task. The primary dependent variable was the number of advantageous choices on each block of 10 trials.

**Main study.** Task procedures for the main study were identical to those used in the pilot study. However, only four blocks of ten trials were administered for each task in the main study. In the pilot study, there was no significant difference in performance on blocks 4 versus 5 in any of the three versions. It was therefore deemed appropriate to reduce the number of trial blocks.

**Results**

**Pilot Study**

Preliminary analyses indicated that there were no significant differences in the performance of children across the two binary-relational versions of the task, $F(1, 23) = .73, p = .401, \eta^2 = .031$. Consequently, subsequent analyses were conducted on a combined binary-relational score that was averaged across the binary-gain and binary-loss versions. The order in which the three versions were presented also had no effect, so this variable was excluded from further analysis.

Number of choices from the advantageous deck were subjected to a 3 (age: 3, 4, 5 years) × 2 (gender: male, female) × 2 (complexity: binary, ternary) × 5 (blocks 1-5) mixed model ANOVA. The analysis indicated significant main effects of complexity, $F(1, 66) = 41.47, p < .001, \eta^2 = .386$, block, $F(4, 264) = 17.69, p < .001, \eta^2 = .211$, and age, $F(2, 66) = 7.99, p < .001, \eta^2 = .195$. Significant 2-way interactions were evident for Complexity × Age, $F(2, 66) = 17.43, p < .001, \eta^2 = .346$, Block × Age, $F(8, 264) = 5.91, p < .001, \eta^2 = .152$, and Complexity × Block, $F(4, 264) = 18.75, p < .001, \eta^2 = .221$. Finally, two significant 3-way interactions were found for Complexity × Blocks × Age, $F(8, 264) = 5.19, p < .001, \eta^2 = .136$, and Complexity × Age × Gender, $F(2, 66) =$
3.24, \( p = .045, \eta^2 = .09 \). The 4-way interaction was not significant, \( F(8, 264) = 1.02, .378, \eta^2 = .03 \).

The Complexity × Blocks × Age interaction was partitioned by complexity. For the combined binary-relational versions, ANOVA revealed significant effects of age, \( F(2, 69) = 4.14, p = .020, \eta^2 = .972 \), indicating that the number of advantageous choices increased with age, and of block, \( F(4,276) = 54.14, p < .001, \eta^2 = .440 \), indicating an increase in advantageous choices across blocks. The Blocks × Age interaction was not significant, \( F(4,276) = 1.26, p = .264, \eta^2 = .035 \). Means and standard errors for the binary-relational versions across the five trial blocks are presented in Figure 4.2a.

Figure 4.2. Mean number (standard error) of advantageous choices on the binary-relational decks (a) and ternary-relational deck (b) of the pilot study (\( N = 72 \)).
For the ternary-relational version there was a significant main effect of block, $F(4, 276) = 5.00, p = .001$, $\eta^2 = .068$, a significant main effect of age, $F(2, 66) = 14.70$, $p < .001$, $\eta^2 = .299$, and a significant Block $\times$ Age interaction, $F(8, 276) = 7.00, p < .001$, $\eta^2 = .168$. Means and standard errors for the interaction are presented in Figure 4.2b. Analysis of simple effects indicated that choices from advantageous decks decreased significantly across blocks for 3-year-olds, $F(4, 92) = 5.11, p = .001$, $\eta^2 = .182$, and increased significantly across blocks for 5-year-olds, $F(4, 92) = 22.24, p < .001$, $\eta^2 =$. No significant effect of trial blocks was evident for the 4-year-olds, $F(4, 92) = 1.29, p = .279$, $\eta^2 = .053$.

Single sample $t$-tests were used to compare the means for each age group and trial block to chance level (5). For the binary-relational versions, the mean scores of 3-, 4- and 5-year-olds for the five trial blocks were all significantly above chance, smallest $t(23) = 2.66, p = .014$. For the ternary-relational version, the 3-year-olds’ mean scores were significantly lower than chance (i.e., fewer advantageous choices) for blocks 2, 3 and 4 ($p$’s < .05). The 5-year-olds’ means were significantly higher than chance (i.e., more advantageous choices) for blocks 2, and 3 ($p$’s < .05) and blocks 4 and 5 ($p$’s < .001). The performance of the 4-year-olds hovered around chance levels.

The Complexity $\times$ Age $\times$ Gender interaction was also partitioned according to complexity (see Figure 4.3). No main effect or interaction involving gender were recorded for the binary-relational versions, however, there was a significant effect of age, $F(2,66) = 4.06, p = .022$, $\eta^2 = .109$. A Scheffe test revealed significant differences between means for the 3-year-olds and 4-year-olds only ($p = .023$). Analysis of the ternary-relational version indicated a significant Gender $\times$ Age interaction, $F(2, 66) = 3.77, p = .028$, $\eta^2 = .102$, and a significant effect of age, $F(2, 66) = 18.14, p < .001$, $\eta^2 = .355$. Means and standard errors for the 2-way interaction are presented in Figure 4.3b.
Analysis of simple effects indicated that 3-year-old females made significantly more advantageous choices than 3-year-old males, $F(1, 23) = 5.14, p = .034, \eta^2 = .189$. There were no differences between genders for the 4-year-olds, $F(1, 23) = 1.18, p = .288, \eta^2 = .051$, or 5-year-olds, $F(1, 23) = .11, p = .741, \eta^2 = .005$.

![Graph](https://via.placeholder.com/150)

**Figure 4.3.** Mean number (standard error) of advantageous choices on the binary-relational versions (a) and the ternary-relational version (b) of the pilot study ($N = 72$).

Performances of the individual children were evaluated against chance. The probability of a correct response on each item was 0.50, thus scores of 32 or more out of 50 are significantly above chance according to the binomial table ($p < .05$). As can be seen in Table 4.1, the majority of children in all age groups performed above chance.
level on the binary-relational versions. In contrast, a minority of 3- and 4-year-olds and a majority of 5-year-olds performed significantly above chance on the ternary-relational version. The median age of attainment corresponds to the age at which 50% of children pass the tasks. The pass-fail data indicated that the age of attainment was three years (or earlier) for the binary-relational versions, but not until five years for the ternary-relational version of the Gambling Task.

Table 4.1

*Frequency of Children in Each Age-Group Performing above Chance Levels on the Combined Binary-Relational and Ternary-Relational Versions in the Pilot Study (N = 72)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Versions</th>
<th>Ternary Relational Version</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>4-years</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>5-years</td>
<td>5</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 4.2 relates the pass-fail scores for the binary-and ternary-relational versions. Fifty-seven percent of children (41) performed consistently, failing or passing versions at both complexity levels. The remaining (31) children showed an inconsistent pattern (failing one level while passing the other). These children were significantly more likely to pass the binary-relational versions and fail the ternary-relational Children’s Gambling Task (83.87%) than to show the reverse pattern (16.13%), McNemar test, $\chi^2(N=72, 1) = 12.90, p < .001$. Thus, the binary- and ternary-relational versions were sufficiently difficult for the age range of children included in this study.
Table 4.2

*Crosstabulations of Children Passing and Failing the Combined Binary- and Ternary-Relational Versions of the Gambling Task in the Pilot Study (N = 72)*

<table>
<thead>
<tr>
<th></th>
<th>Ternary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>Fail</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Pass</td>
<td>26</td>
<td>28</td>
</tr>
</tbody>
</table>

**Main Study**

For the main study, the distribution of the 5-year-old children on the ternary-relational version was significantly skewed to a moderate degree. Inspection of the raw scores indicated that this was due to a single outlying score. This data point represented a 5-year-old child who had scored zero on the ternary relational deck. This score was removed, and the analyses re-run. The removal of this score did not impact on significance of the findings and thus all data-points were retained in the final analysis.

Preliminary analyses indicated that there were no significant differences in the performance of children across the two binary-relational versions, \( F(1,119) = 1.11, p = .295, \eta^2 = .009 \). Consequently, subsequent analyses were conducted on a combined binary-relational score that was averaged across the binary-gain and binary-loss versions. The order in which the three versions were presented also had no effect so this variable was excluded from further analysis, \( F(3, 88) = 0.42, p = .736, \eta^2 = .012 \).

A 4 (age: 3, 4, 5, 6 years) × 2 (gender: male, female) × 2 (complexity: binary, ternary) × 4 (blocks: 1-4) mixed ANOVA was first conducted. The analysis indicated
significant main effects of complexity, $F(1, 112) = 128.74, p < .001, \eta^2 = .535$, block,
$F(3, 336) = 99.30, p < .001, \eta^2 = .470$, age, $F(3, 112) = 33.78, p < .001, \eta^2 = .475$, and
gender, $F(1, 112) = 4.18, p = .043, \eta^2 = .036$, with females outperforming males.
Significant 2-way interactions were evident for Complexity $\times$ Age, $F(3, 112) = 24.71, p < .001, \eta^2 = .398$, and Block $\times$ Age, $F(9, 336) = 12.25, p < .001, \eta^2 = .247$. Finally, a
significant 3-way interaction was found for Complexity $\times$ Blocks $\times$ Age, $F(9, 336) = 5.17, p < .001, \eta^2 = .122$.

The 3-way interaction was partitioned according to complexity (see Figure 4.4).
For the combined binary-relational score, ANOVA revealed a significant effect of
block, $F(3, 348) = 182.32, p < .001, \eta^2 = .611$, indicating an increase in advantageous
choices across blocks of trials. No significant main effect of age, $F(3, 116) = 1.73, p = .146, \eta^2 = .042$, or Block $\times$ Age interaction, $F(3, 166) = 2.00, p = .117, \eta^2 = .055$, was
found. Means and standard errors for the performance of the three age groups across the
five trial blocks are presented in Figure 4.4a.
For the ternary-relational version there was a significant main effect of block,
$F(3, 348) = 17.12, p < .001, \eta^2 = .129$, a significant main effect of age, $F(3, 116) = 34.15, p < .001, \eta^2 = .469$, and a significant Blocks $\times$ Age interaction, $F(3, 116) = 9.94, p < .001, \eta^2 = .205$. Means and standard errors for the interaction are presented in
Figure 4.4b. Analysis of simple effects indicated that choices from advantageous deck
increased significantly across blocks for 5-year-olds, $F(3, 87) = 49.26, p < .001, \eta^2 = .629$, and 6-year-olds, $F(3, 87) = 45.64, p < .001, \eta^2 = .611$. No significant effect of trial
blocks was evident for the 3-year-olds, $F(3, 87) = 1.76, p = .160, \eta^2 = .057$, or 4-year-olds, $F(3, 87) = .52, p = .672, \eta^2 = .018$. 
Figure 4.4. Mean number (standard error) of advantageous choices on the combined binary-relational versions (a) and the ternary-relational version (b) across the blocks of 10 trials in the main study ($N = 120$).

Single sample $t$-tests showed that for the binary-relational versions, the mean scores of 3-, 4-, 5-, and 6-year-olds for the four trial blocks were all significantly above chance level (5), smallest $t(29) = 5.46, p < .001$. For the ternary-relational version, the 3-year-olds’ mean scores were significantly lower than chance (i.e., fewer advantageous choices) for blocks 2, 3 and 4 ($p$’s < .05). The 5-year-olds’ means were significantly higher than chance (i.e., more advantageous choices) for blocks 2, 3, and 4 ($p$’s < .001).
The 6-year-olds’ mean were significantly higher than chance across all trail blocks \((p < .001)\). The performance of the 4-year-olds hovered around chance levels.

**Total Score.** The Age \(\times\) Complexity interaction was re-examined using a total score that was combined across the trial blocks (maximum = 40), in order to allow for comparability with the performance of tasks in later chapters. The analysis indicated significant main effects of complexity, \(F(1, 116) = 125.85, p < .001, \eta^2 = .520\), and age, \(F(3, 116) = 33.31, p < .001, \eta^2 = .463\), and a significant Complexity \(\times\) Age interaction, \(F(3, 116) = 24.08, p < .001, \eta^2 = .384\). Means and standard errors for the interaction are presented in Figure 4.5. Analysis of simple effects indicated that there was no significant difference in the performance of the four age groups on the binary-relational items, \(F(3, 116) = 1.83, p = .146, \eta^2 = .045\). In contrast, the performance of children on the ternary-relational items increased significantly across the four age groups, \(F(3, 116) = 34.15, p < .001, \eta^2 = .469\). A Scheffe test revealed significant differences between means for 3-year-olds and 4-, 5- and 6-year-olds \((p’s < .001)\); 4-year-olds and 5- and 6-year-olds \((p’s < .001)\).

![Figure 4.5. Mean number of correct choices for each age group on the binary- and ternary-relational items of the Gambling task, pilot study \((N = 120)\).](image-url)
Performances of the individual children were evaluated against chance. The probability of a correct response on each item was 0.50, thus scores of 25 or more out of 40 are significantly above chance according to the binomial table \( p < .05 \). As can be seen in Table 4.3, the majority of children in all age groups performed above chance level on the binary-relational versions. In contrast, a minority of 3- and 4-year-olds and a majority of 5-year-old and 6-year-olds performed significantly above chance on the ternary-relational version. The median age of attainment was three years (or earlier) for the binary-relational versions, but not until five years for the ternary-relational version of the Gambling Task.

Table 4.3

Frequency of Children in Each Age-Group Performing above Chance Levels on the Combined Binary- and Ternary-Relational Versions in the Main Study \((N = 120)\)

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Decks</th>
<th>Ternary Relational Deck</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>4-years</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>5-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>6-years</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 4.4 relates the pass-fail scores for the binary- and ternary-relational versions. Sixty-five percent of children \((78)\) performed consistently, failing or passing versions at both levels of complexity. The remaining \((42)\) children showed an inconsistent pattern (failing one level while passing the other). These children were
significantly more likely to pass the binary-relational versions and fail the ternary relation version of the Children’s Gambling task (97.62%) than to show the reverse pattern (2.38%). McNemar test, $\chi^2(N = 120, 1) = 40.02, p < .0001$.

Table 4.4
Crosstabulation of Children Passing and Failing the Combined Binary- and Ternary-Relational Versions in the Main Study ($N = 120$)

```
           Ternary
           Fail   Pass
   Binary
   Fail    3     1
   Pass    41    75
```

Discussion

The results of both the pilot data and main study indicate significant development of complex decision-making skills during the preschool years. Consistent with Kerr and Zelazo (2004), 3-year-olds made more disadvantageous choices across the trial blocks of the ternary-relational deck than would be expected by chance. Four-year-old children made more advantageous choices than 3-year-olds, however, their performance hovered at chance levels only. Five-year-old and 6-year-old children made more advantageous choices than would be expected by chance across the majority of trial blocks. Thus, the original version of the Children’s Gambling Task was not passed until five years of age.

In contrast, performance of all age groups on the newly created binary-relational versions of the task was above chance level across all trial blocks. The successful completion of both binary-relational versions demonstrates that young children are able
to perform all components of the gambling task. They are able to process differences between decks in gains alone on the binary-gain version and differences in losses alone on the binary-loss version. What they cannot do is integrate these differences about both gains and losses in order to identify the advantageous deck in the ternary-relational version. Thus, consistent with relational complexity theory 3- to 6-year-olds passed the binary relational items, while the ternary-relational items were not passed until five years of age.

The results can also be interpreted in terms of the CCC theory. According to this approach, younger children may experience difficulty on the original version of the Children’s Gambling Task due to an inability to formulate a higher order rule that allows them coordinate initial discriminations between the two decks that is based on gains, with emerging evidence regarding losses. In contrast, older children are better able to reflect on these conflicting discriminations and formulate a higher-order rule that allows them to appreciate net gains (Kerr & Zelazo, 2004). The less complex versions of the task, however, manipulated the magnitude of the gains or losses independently of the other. As such, these versions of the task do not require the formulation of a higher order relation, and therefore these items posed no difficulty to children in any of the age groups examined.

Kerr and Zelazo (2004) further suggested that functional immaturity of the OFC in 3-year-olds may result in a failure to develop somatic markers associated with the disadvantageous decks in the Children’s Gambling Tasks. However, the current findings counter this suggestion because 3-years-old succeeded on the less complex binary-relational versions. Thus, the Somatic Marker Hypothesis may either not be entirely applicable, or may need to be adjusted to incorporate consideration of the complexity of the tasks involved. Further research could determine whether patients with OFC lesions succeed when task complexity is reduced.
A somewhat curious finding in the pilot study only was that while the majority of the 4-year-olds performed above chance levels on the binary-relational versions, 10 of them did not. This finding may simply be a characteristic of the particular sample of 4-year-olds used in the pilot study, as the pattern was not seen in the main study.

The results of the current research are inconsistent with some previous research that has indicated gender effects in the performance of tasks associated with the OFC. The current study provided evidence of female superiority in the three-year-old sample in the pilot data, and across all age groups in the main study. Consistent with this, Garon and Moore (2004) found that females chose significantly more cards from the advantageous decks than would be expected by chance by the second block, whereas males demonstrated no significant difference in their choice of decks. In contrast, Kerr and Zelazo (2004) found preliminary evidence of male superiority for the 3-year-olds on their version of the Gambling Task. Clearly, further research is needed to disentangle any possible influence of gender across age groups in the performance of these complex decision-making tasks.

In summary, both pilot data and main study indicated that 3-, 4-, 5- and 6-year-olds made advantageous card selections on the binary-relational versions, but only 5- and 6-year-olds did so on the ternary-relational version. Thus, the failure of the younger age groups on the ternary-relational items was not due to an inability to process any individual component of the task, such as recognizing disadvantageous items, but rather to an inability to integrate information about both gains and losses across all items. These results indicate that the primary requirement for successful completion of the original version of the Children’s Gambling Task is the children’s ability to process complex relations.
CHAPTER 5. FUTURE-ORIENTED DECISION-MAKING

The ability to think about the future is an important aspect of human cognition and necessary for adaptive functioning (Atance & O’Neill, 2001). Investigations into this distinctly human behaviour first focused on the ability of individuals to re-experience the past. To this end, Tulving (1985) differentiated between two aspects of memory. Episodic memory was described as the system that allows an individual to remember personally experienced events, and to travel back in time to re-experience them. In contrast, semantic memory was described as an individual’s knowledge of the world. Drawing upon this distinction, Atance and O’Neill (2001) contrasted episodic future thinking with semantic future thinking. Episodic future thinking was defined as the ability to project the self into the future to pre-experience an event. This was compared to an individual’s more general knowledge about the future, labeled semantic future thinking. After reviewing relevant cognitive, social and neuropsychological literature, Atance and O’Neill (2001) concluded that the frontal cortex may play an important role in a wider neuronal network responsible for control of episodic future thinking. Further, the emergence of the planning and anticipatory behaviours required by episodic future thinking may not occur until four to five years of age.

Busby and Suddendorf (2005) examined the ability of 3-, 4- and 5-year-old children to mentally time travel into the past (episodic memory) and future (episodic future memory). The children were asked to report something they did yesterday, and something that they were going to do tomorrow. The majority of 4-year-olds were able to correctly report events they had experienced yesterday (55% in Study 1; 56% in Study 2) and would experience tomorrow (65% and 69%). While 3-year-olds were able to report events in response to the questions, the events they reported were often incorrect. Only a minority of 3-year-olds were able to correctly report events they had experienced yesterday (30% in Study 1; 25% in Study 2) and would experience
tomorrow (30% and 31%). Five-year-old children were also included in Study 2, with results indicating that 75% of 5-year-olds performed correctly on the yesterday question and 63% passed the tomorrow question. The authors concluded that the incorrect reports by 3-year-olds reflected their inability to mentally travel in time. As such, the emergence of episodic memory and episodic future memory may occur in tandem at four to five years of age (Busby & Suddendorf, 2005).

The emergence of episodic future thinking in the preschool years is consistent with research into the development of theory-of-mind abilities. Such research has found consistent support for the emergence of the ability to attribute mental states to others at approximately 4.5 years of age (Wellman, Cross & Watson, 2001). Thus, it has been proposed that projecting one’s own future mental states as required by episodic future thinking, might be just as difficult, and might even be based on similar mechanisms, as the ability to imagine the mental state of others (Suddendorf & Corbalis, 1997).

Suddendorf and Busby (2005) created a non-verbal test of mental time travel that provided participants with the opportunity to anticipate a future need such as avoiding boredom. In this ‘rooms’ task children spent time in an empty room which contained only a puzzle board without the puzzle pieces, and an active room which contained a range of different toys including the puzzle pieces. Children were told they could take one of the items from the active room to the empty room. A control group of children were exposed to the same procedures as the experimental group, however, the puzzle board was not present in the empty room. This preliminary study found that the toy selections of 3-year-olds were not influenced by whether they had encountered the puzzle board in the empty room. In contrast, the toy selections of 4- and 5-year-old children were dependent on what they had experienced in the empty room. That is, the older children might have imagined themselves back in the empty room with the puzzle board and anticipated the need to have the puzzle pieces in order to be able to play the game.
The ability of children to consider the future consequences of their current behaviour has also been examined using traditional delay-of-gratification paradigms. This research has examined children’s ability to wait for a more desirable, but delayed, outcome rather than accept a less desirable, but immediate, outcome (Mischel & Mischel, 1983). In such research, the children are presented with two incentives (e.g., food treats) and are asked to choose which one they like best. They are then told that to obtain the more preferred incentive that they must wait alone until the researcher returns (usually 15 minutes). The children can stop the delay at any time by ringing a bell, at which point the experimenter will return immediately; however, the child will lose the more preferred incentive and receive the less preferred one. Using this paradigm, Mischel and Mischel (1983) found that children three to four years of age had difficulty in delaying their gratification, as measured by how long they would wait for the preferred outcome.

Thompson, Barresi and Moore (1997) modified the delay-of-gratification task in order to assess prudence and altruism in situations involving future desires. The research placed 3- to 5-year-old children in conflict situations where they were required to make a choice between two desirable alternatives involving stickers. Three altruism tasks were presented whereby children were given a choice between: (a) an immediate reward (1 sticker for self now) and an immediate reward for self and other (1 sticker each now); (b) an immediate reward (2 stickers for self now) and a reward for self and other (1 sticker each now); and (c) an immediate reward (2 stickers for self now) and a shared delayed reward (2 stickers each later). A prudence task was included in which children were given a choice between an immediate reward (1 sticker for self now) and a larger delayed reward (2 stickers for self later). While the previous delay-of-gratification paradigm used the length of delay as their outcome measure, the dependent variable in Thompson et al.’s (1997) study was the number of times children chose the
delayed option. In trials with no delay, Thompson et al. (1997) found that children of all ages chose to share, suggesting that the basic capacity to engage in prosocial behaviour is in place by three years of age. In the trials involving delay, however, the results indicated that when given the option to choose between a small reward now or a large reward later (prudence) and between a reward for self now or a reward for self and other later (altruism), 4- to 5-year-old children demonstrated significantly more prudence and altruism than 3-year-old children. Furthermore, they found that future oriented prudence and altruism were significantly correlated in 3-year-olds, suggesting that being able to consider future scenarios for self and other were closely related.

The altruism and prudence items were examined further by Moore, Barresi and Thompson (1998) who also included a theory-of-mind task which required children to consider the mental states of themselves and others, and an inhibition task in which children had to inhibit pointing to a baited box in order to win the cookie inside. In addition to replicating the previous findings, the results showed that for 4-year-olds, future oriented altruism was related to performance on measures of theory-of-mind. For 3-year-olds future-oriented prudence was significantly correlated with the inhibition task. The authors concluded that future-oriented prudence and altruism are developmentally linked to the ability to inhibit a salient response and the ability to imagine conflicting mental states.

Principe and Zelazo (2005) adapted the procedures of Thompson et al. (1997) to include additional trial types. Nine trial types were administered by using three types of rewards (stickers, pennies, candies) and three types of choice (one now vs. two later, one now vs. four later, one now vs. six later). The participants included 3- and 4-year-old children, half of whom chose for them self (self condition) and half who chose for the experimenter (other condition). The results indicated that 3-year-olds typically chose impulsively for themselves, however, they usually chose to delay for the experimenter.
In comparison, the 4-year-olds chose to delay more often the 3-year-olds in the self condition but less often than 3-year-olds in the other condition. The authors proposed that 3-year-olds were able to take either a first person perspective (for self) or a third person perspective (for other). It is not until four years of age that children are able to integrate these perspectives in order to consider a third-person perspective in the self condition and the experimenter’s perspective in the other condition.

The development of delay of gratification is suggested to be particularly important for social and personality development. Even as young as two years of age, children can respond altruistically to another’ emotional distress, for example by giving a favourite toy away to a distressed person (Zahn-Waxler, Radke-Yarrow & Wagner, 1992). Longitudinal research has also indicated that the ability to delay gratification during preschool predicts patterns of adaptive functioning into adulthood. For example, Shoda, Mischel and Peake (1990) demonstrated that preschool children who waited longer were perceived by their parents to be more cognitively and socially competent, and better able to cope with stress during adolescence. In order to deal effectively with social interactions, children are also often required to understand and deal with possible future situations (Thompson et al., 1997). Thus, the development of future-oriented prudence and altruism is an important developmental milestone, and one that warrants further attention in relation to the development of hot executive functions.

Relation Of Future-Oriented Decision-Making to the OFC

The delay of gratification paradigm is considered to be a measure of hot executive function, as the task involves consideration of the arousing qualities of the preferred reward (Zelazo & Müller, 2002). As aforementioned, the OFC has a specific role in processing the reward values of stimuli. Significant activation of the OFC has been found during tasks that require not only the monitoring of reward values, but also the holding in mind of reward values of past and future stimuli and responses (Elliot et
According to Metcalfe and Mischel’s (1999) hot/cool system of willpower, the object of desire in the delay-of-gratification task is “hotly” displayed as the rewards are directly and saliently exposed to the child’s attention. Thus, a significant involvement of the OFC is anticipated during the performance of this task. However, Metcalfe and Mischel (1999) also suggested that episodic representation and thought may be a primary function of the cool, cognitive know system. Thus, if the delay of the gratification paradigm is represented as involving episodic future thinking, it raises questions regarding the ‘hotness’ of this task.

**Complexity Analysis of Future-Oriented Decision-Making Tasks**

The current study aimed to modify the prudence/altruism task developed by Thompson et al. (1997) in order to incorporate a complexity manipulation. The original task incorporated the following four types of items, each with two choices:

1. Shared gratification without cost: (a) an immediate reward (1 sticker for self now) and (b) an immediate reward for self and other (1 sticker each now).
2. Shared gratification with cost: (a) a immediate reward (2 stickers for self now) and (b) a reward for self and other (1 sticker each now).
3. Delay of self gratification: (a) an immediate reward (1 sticker for self now) and (b) a larger delayed reward (2 stickers for self later).
4. Delay of shared gratification: (a) an immediate reward (2 stickers for self now) and (b) a shared delayed reward (2 stickers each later).

According to the relational complexity approach, Items 3 and 4 of this task involve ternary relations. This is because the binary relation between the two magnitudes of reward, as indicated in options (a) and (b), is modified by the delay variable. Thus, Items 3 and 4 involve three interacting variables (magnitude a, magnitude b, delay). Item 2 also involves a ternary relation, as the binary relation between the two magnitudes of the reward is influenced by a third variable, cost to the
self. Thus, this item also results in three interacting variables (magnitude a, magnitude b, cost). However, task complexity is lower in Item 1 because the delay and cost variables are omitted, leaving only the two magnitudes of reward variables to be considered. Hence, Item 1 is binary-relational.

The current research used these existing task procedures to model binary- and ternary-relational items. The study included three types of ternary-relational trials (see Table 5.1). The Self Gratification, and Shared Gratification trials correspond to items 3 and 4 described above. The additional trial type, Other Gratification, is similar to type 4 except that the options refer to other rather than self. As can be seen in the first column, the ternary-relational items incorporate three variables (magnitude a, magnitude b, delay). To succeed on these items, children would have to understand that a larger reward would be obtained by choosing the delayed option. The equivalent binary-relational items involve two magnitudes of rewards only (magnitude a, magnitude b), as the delay variable has been removed.

Table 5.1

*Example Binary- and Ternary-Relational Items in the Future-Oriented Decision Making Task*

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Choices for Binary Items</th>
<th>Choices for Ternary Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared gratification</td>
<td>(a) 1 for self now vs</td>
<td>(a) 1 sticker for self now vs</td>
</tr>
<tr>
<td></td>
<td>(b) 1 each now</td>
<td>(b) 1 sticker each later</td>
</tr>
<tr>
<td>Self Gratification</td>
<td>(a) 1 for self now vs</td>
<td>(a) 1 sticker for self now vs</td>
</tr>
<tr>
<td></td>
<td>(b) 2 for self now</td>
<td>(b) 2 stickers for self later</td>
</tr>
<tr>
<td>Other Gratification</td>
<td>(a) 1 for other now vs</td>
<td>(a) 1 sticker for other now vs</td>
</tr>
<tr>
<td></td>
<td>(b) 2 for other now</td>
<td>(b) 2 stickers for other later</td>
</tr>
</tbody>
</table>
Consistent with the relational complexity approach and the predictions outlined in Chapter 3, it is anticipated that 3-, 4-, 5- and 6-year-old children would succeed on the binary-relational items of the future-oriented decision-making task. The ternary-relational items included a delay, thereby requiring children to project into the future to pre-experience an event. Delay constitutes an additional variable, which means that complexity is increased relative to the binary-relational items. It was expected that 3- and 4-year-old children would experience difficulty incorporating the three variables required for success on these ternary-relational items. On the other hand, 5- and 6-year-olds should be able to process the three variables and succeed on the ternary-relational items of the future-oriented decision-making task.

Method

Materials.

For every child tested, two sticker books were utilized: one for the child and one for the researcher (who began a new book with each child tested). Numerous sheets of stickers were used as stimuli in the trials (with approximately 30 stickers used per child), and two envelopes were used to contain delayed-reward stickers (one envelope each for the child and the researcher). The response sheets used for recording each child’s selections are presented in Appendix C.

Procedure

Pilot study. The task began with the experimenter sitting directly across from the child with a sticker book placed in front of each of them. The child was encouraged to write their name on the front of the book (the names of younger children were written by the experimenter) and they were told that the sticker book was theirs to keep. The children were introduced to the task with the experimenter explaining that this game involved both the experimenter and the child receiving stickers to place in their
respective sticker books. The children were told that they were allowed to choose whether they alone would receive a sticker, or whether the experimenter would get stickers too. It was explained that the child could choose how many stickers each person would get, and whether they would have their stickers right away to put in their sticker-books or whether they would wait and get their stickers at the end of the game.

Envelopes for delayed reward stickers were shown to each child, and they were assured that, if delayed-reward options were chosen, the stickers would be placed in the envelopes and saved until the end of the game. It was also explained that, at the end of the game, all the stickers would be taken out of the envelopes and returned to the child and the research assistant so that they could put them into their sticker books and take them home.

Each child received four binary-relational items and four ternary-relational items of each of the trial types (shared gratification, self gratification, other gratification). Thus, 12 items were presented at each level of complexity. The numerosity of the stickers was varied across trials, ranging from one to four. The 24 items (12 binary-relational and 12 ternary-relational) are displayed in Table 5.2 and were presented in a counter-balanced order.

According to the child’s expressed preference, the stickers were either given to them to be placed in their sticker book immediately, or they were placed in the envelope and kept out of the child’s sight until the end of the game. The experimenter responded in a uniform, mildly positive manner to any choice made by the child. Upon completion of the 24 trials, the delayed reward stickers were given to the child to be placed in their sticker books. Responses were scored as correct if the option with the highest magnitude of reward was chosen. This corresponded to the delayed option on the ternary-relational items. Individual scores (out of 12) were calculated for each level of complexity.
Table 5.2

*Binary- and Ternary-Relational Items in the Future-Oriented Decision-Making Task*

*(correct responses are highlighted)*

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Choices for Binary-Relational Items</th>
<th>Choices Ternary-Relational Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Option 1</td>
<td>Option 2</td>
</tr>
<tr>
<td>Shared Gratification</td>
<td>4 each now</td>
<td>4 for self now</td>
</tr>
<tr>
<td></td>
<td>3 for self now</td>
<td>3 each now</td>
</tr>
<tr>
<td></td>
<td>2 for self now</td>
<td>2 each now</td>
</tr>
<tr>
<td></td>
<td>1 each now</td>
<td>1 for self now</td>
</tr>
<tr>
<td>Self gratification</td>
<td>4 for self now</td>
<td>2 for self now</td>
</tr>
<tr>
<td></td>
<td>2 for self now</td>
<td>4 for self now</td>
</tr>
<tr>
<td></td>
<td>2 for self now</td>
<td>1 for self now</td>
</tr>
<tr>
<td></td>
<td>1 for self now</td>
<td>2 for self now</td>
</tr>
<tr>
<td>Other gratification</td>
<td>2 for other now</td>
<td>1 for other now</td>
</tr>
<tr>
<td></td>
<td>2 for other now</td>
<td>4 for other now</td>
</tr>
<tr>
<td></td>
<td>2 for other now</td>
<td>1 for other now</td>
</tr>
<tr>
<td></td>
<td>2 for other now</td>
<td>4 for other now</td>
</tr>
</tbody>
</table>

*Main study.* The administration of the future-oriented decision-making task in the main study was identical to that used in the pilot study.

*Results*

*Pilot Study*

The order in which the two levels of complexity were presented had no effect so this variable was excluded from further analysis, $F(1, 66) = 2.73, p = .104, \eta^2 = .043$. A 3 (age: 3, 4, 5 years) × 2 (gender: male, female) × 2 (complexity: binary, ternary) × 3
Development of Executive Functions

(trial type: shared gratification, self gratification, other gratification) mixed ANOVA was conducted. The analysis indicated significant main effects of complexity, \( F(1, 66) = 123.85, p < .001, \eta^2 = .652 \), and age, \( F(2, 66) = 20.43, p < .001, \eta^2 = .382 \). Three significant 2-way interactions were also evident.

Firstly, a significant Complexity × Age interaction was found, \( F(2, 66) = 9.23, p < .001, \eta^2 = .219 \). Means and standard errors for the interaction are presented in Figure 5.1. Analysis of simple effects indicated that there was no significant difference in the performance of the three age groups on the binary-relational items, \( F(2, 69) = 2.95, p = .059, \eta^2 = .079 \). In contrast, the performance of children on the ternary-relational items increased significantly across the three age groups, \( F(2, 69) = 24.98, p < .001, \eta^2 = .844 \). A Scheffe test revealed significant differences between means for 3-year-olds and 5-year-olds \((p < .001)\); and 4-year-olds and 5-year-olds \((p < .001)\).

\[\text{Figure 5.1. Mean number of correct choices for each age group on the binary- and ternary-relational items of the future-oriented decision-making task, pilot study (}\ N = 72\text{)}\]

A significant Trial type × Complexity interaction was also observed, \( F(2, 132) = 6.80, p = .002, \eta^2 = .093 \). Means and standard errors for the interaction are presented in
Figure 5.2. Analysis of simple effects indicated that the effect of trial type was significant for the binary-relational items, $F(2, 142) = 7.09, p = .001, \eta^2 = .091$. Contrast analysis indicated that the mean for self gratification was significant higher than for other gratification, $F(6, 66) = 11.64, p = .001, \eta^2 = .141$, and that shared gratification was significantly higher than for other gratification, $F(6, 66) = 7.32, p = .009, \eta^2 = .093$. The effect of trial type was not significant for the ternary-relational items, $F(2, 142) = 2.15, p = .120, \eta^2 = .029$.

Figure 5.2. Mean number of correct responses across the three item types and two levels of complexity in the pilot study (N = 72).

A significant Trial type $\times$ Gender interaction was also observed, $F(2, 132) = 6.74, p = .002, \eta^2 = .093$, presented in Figure 5.3. Analysis of simple effects indicated there was no significant difference in the performance of males across the three trial types, $F(2, 70) = 0.57, p = .568, \eta^2 = .016$. A significant difference was recorded in the performance of females across the three trial types, $F(2, 69) = 9.08, p < .001, \eta^2 = .206$. Contrast analysis indicated significant differences between the performance of females on the shared gratification and self gratification items, $F(3, 33) = , p < .001, \eta^2 = .206$. Girls performed better on the shared gratification items than the self gratification items.
Single sample $t$-tests were used to compare the means for each age group to their respective chance levels (6). For the binary-relational items, the mean scores of 3-, 4-, and 5-year-olds were all significantly above chance, smallest $t(23) = 5.52$, $p < .001$. For the ternary-relational items, the mean scores were significantly below chance for 3-year-olds, $t(23) = -5.99$, $p < .001$, and 4-year-olds, $t(23) = -2.38$, $p = .026$, thus indicating a bias towards selecting the immediate reward option. The mean score of 5-year-olds was significantly higher than chance, $t(23) = 4.26$, $p < .001$ (see Figure 5.1).

Performances of the individual children were evaluated against chance. The probability of a correct response on each item was 0.50, thus scores of 9 or more out of 12 are significantly above chance according to the binomial table ($p < .05$). As can be seen in Table 5.3, the majority of children in all age groups performed above chance level on the binary-relational items. A minority of 3- and 4-year-olds and a majority of 5-year-olds performed significantly above chance on the ternary-relational items. The pass-fail data indicated that the median age of attainment was three years (or earlier) for the binary-relational items, but not until five years for the ternary-relational items of the future-oriented decision-making task.
Table 5.3

*Frequency of Children Performing above Chance Levels on the Two Levels of Complexity of the Future-Oriented Decision-Making Task, Pilot Study (N = 72)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Items</th>
<th>Ternary Relational Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>4-years</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>5-years</td>
<td>3</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 5.4 relates the pass-fail scores for the binary- and ternary-relational items. Fifty-three percent of children (38) performed consistently, failing or passing both complexity levels. The remaining (34) children showed an inconsistent pattern (failing one level, passing the other). These children all passed the binary-relational items and failed the ternary-relation items, McNemar test, $\chi^2(N = 72, 1) = 32.03, p < .001$.

Table 5.4

*Crosstabulation of Children Passing and Failing the Binary- and Ternary-Relational Items of the Future-Oriented Decision-Making Task in the Pilot Study (N = 72)*

<table>
<thead>
<tr>
<th>Binary</th>
<th>Ternary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
</tr>
<tr>
<td>Fail</td>
<td>21</td>
</tr>
<tr>
<td>Pass</td>
<td>34</td>
</tr>
</tbody>
</table>
Main Study

The order in which the two levels of complexity were presented had no effect so this variable was excluded from further analysis, $F(1, 104) = 2.23, p = .139, \eta^2 = .021$. The number of correct choices children made was subjected to a 4 (age: 3, 4, 5, 6 years) $\times$ 2 (gender: male, female) $\times$ 2 (complexity: binary, ternary) $\times$ 3 (trial type: shared gratification, self-gratification, other gratification) ANOVA. The result indicated significant main effects of complexity, $F(1, 112) = 292.26, p < .001, \eta^2 = .723$, and age, $F(3, 112) = 52.57, p < .001, \eta^2 = 585$. Two significant 2-way interactions were also evident.

Firstly, a significant Complexity $\times$ Age interaction was found, $F(3, 112) = 43.65, p < .001, \eta^2 = .539$. Means and standard errors for the interaction are presented in Figure 5.4. Analysis of simple effects indicated that the performance of children increased significantly across the four age groups on both the binary-relational, $F(3, 116) = 2.71, p = .048, \eta^2 = .065$, and ternary-relational items, $F(3, 116) = 66.14, p < .001, \eta^2 = .631$. The magnitude of this effect however, was much stronger on the ternary-relational items than the binary-relational items. For the binary-relational items, a Scheffe test indicated that there was no significant difference between the means for any of the age groups. For the ternary-relational items, a Scheffe test revealed significant differences between means for: 3-year olds and 4-, 5- and 6-year-olds (smallest $p = .001$); 4-year-olds and the 5- and 6-year-olds ($p$’s < .001).
Figure 5.4. Mean number of correct responses for the four age groups for the binary- and ternary-relational items of the future-oriented decision-making task, main study ($N = 120$).

A significant Trial type $\times$ Complexity interaction was also observed, $F(2, 224) = 9.58, p < .001, \eta^2 = .079$. Means and standard errors for the interaction are presented in Figure 5.5. Analysis of simple effects indicated that the performance of children differed significantly across the trial types for binary-relational items, $F(2, 232) = 11.00, p < .001, \eta^2 = .84$. Contrast analysis indicated that the means for self gratification were significantly higher than for shared gratification, $F(6, 114) = 10.81, p = .001, \eta^2 = .083$, and self gratification was significantly higher than other gratification, $F(6, 114) = 19.29, p < .001, \eta^2 = .139$. In contrast, there was no significant effect of trial type on the ternary-relational items, $F(2, 232) = 2.47, p = .087, \eta^2 = .020$. 
Figure 5.5. Mean number of correct responses for the three item types and the two levels of complexity in the main study ($N = 120$).

Single sample $t$-tests were used to compare the means for each age group to their respective chance levels (6) for the total binary-relational and ternary-relational scores. For the binary-relational items, the mean scores of 3-, 4-, 5-, and 6-year-olds were all significantly above chance, smallest $t(29) = 11.23, p < .001$. For the ternary-relational items, the mean scores were significantly below chance for 3-year-olds, $t(29) = -11.95, p < .001$, and 4-year-olds, $t(29) = -2.10, p = .044$. The mean scores of 5- and 6-year-olds were significantly higher than chance ($p$’s < .001).

Performances of the individual children were evaluated against chance. As can be seen in Table 5.5, the majority of children in all age groups performed above chance level on the binary-relational items. A minority of 3- and 4-year-olds and a majority of 5- and 6-year-olds performed significantly above chance on the ternary-relational items. The pass-fail data indicated that the median age of attainment was five years for the ternary-relational items of the future-oriented decision-making task, and three years (or earlier) for the binary-relational items.
Table 5.5

*Frequency of Children Performing above Chance Levels on the Two levels of Complexity of the Future-Oriented Decision-Making Task in the Main Study (N = 120)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Items</th>
<th></th>
<th>Ternary-Relational Items</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>6</td>
<td>24</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>4-years</td>
<td>3</td>
<td>27</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>5-years</td>
<td>2</td>
<td>28</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>6-years</td>
<td>0</td>
<td>30</td>
<td>7</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 5.6 relates the pass-fail scores for the binary- and ternary-relational items. Forty-six percent of children (55) performed consistently, failing or passing both levels of complexity. The remaining (65) children showed an inconsistent pattern. These children were significantly more likely to pass the binary-relational items and fail the ternary-relational items of the future-oriented decision-making task (96.92%) than to show the reverse pattern (3.17%), McNemar test, \( \chi^2(N = 120, 1) = 55.36, p < .001. \)

Table 5.6

*Crosstabulation of Children Passing and Failing the Binary- and Ternary-Relational Items of the Future-Oriented Decision-Making Task (N = 120).*

<table>
<thead>
<tr>
<th>Binary Fail</th>
<th>Ternary Fail</th>
<th>Ternary Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Pass</td>
<td>63</td>
<td>46</td>
</tr>
</tbody>
</table>
Development of Executive Functions

Discussion

The results of both the pilot data and main study indicate significant development of future-oriented decision-making skills during the preschool years. Consistent with relational complexity theory, the majority of 3-, 4-, 5- and 6-year-olds passed the binary-relational items of this task. The ternary-relational items incorporated a delay variable, and thus involved three interacting variables. The majority of 3- and 4-year-olds failed the ternary-relational items, demonstrating a systematic bias against the delay option in favour of the immediate option. The majority of 5- and 6-year-olds passed the ternary-relational items, selecting the delayed options that had the higher magnitude of reward.

The performance of children on the ternary-relational items is consistent with the findings of Busby and Suddendorf (2005) on their measure of future episodic memory. These authors reported that the emergence of episodic future memory occurred around four to five years of age. Prior to this age, young children were unable to mentally travel in time. As described previously, the delay variable that was incorporated in the ternary-relational items required the ability to mentally travel in time into the future. Thus, the performance of 3- and 4-year-olds in the current research might have also reflected an inability to mentally travel into the future. Without this ability to project oneself into the future and pre-experience the event, young children might have been unable to appreciate the greater magnitude of rewards that could be obtained by choosing the delay variable.

The results of the mean levels of performance on the ternary-relational items are also consistent with the findings of Thompson et al. (1997) on the future-oriented prudence and altruism task. These results indicated that for trials involving delay, 4- to 5-year-old children demonstrated significantly more prudence and altruism than 3-year-old children. Thus, consistent with the performance of 5- and 6-year-olds in the current
sample, these older age groups were able to choose a larger reward for themselves for later (prudence) or to choose a reward for both the self and other for later (altruism). Using an adaptation of this task, Principe and Zelazo (2005) suggested that by four years of age children were able to integrate first and third person perspectives in order to choose the larger response options in the self and other conditions. However, the younger age of attainment found by these authors might be attributable to the larger reward quantities presented in their version of the task (which included up to six rewards) which might have made the differences between the response options more salient for the younger children.

A consistent finding across both the pilot data and main study was the influence of item type on the binary-relational items. Across both sets of results, highest levels of performance were recorded on the self-gratification items, followed by the shared-gratification items, and then other gratification items. Thus, the ability to choose a higher magnitude of reward for themselves may be easier for preschool children than to share with another, or to choose a higher magnitude of reward for the other. It should be noted, however, that despite these differences, 72.22% of 3-year-olds, 77.78% of 4-year-olds, 90.74% of 5-year-olds and 100% of 6-year-olds performed significantly above chance on the binary-relational items (pilot and main studies combined). This result is also consistent with the findings of Thompson et al. (1997) who found that the basic capacity to engage in prosocial behaviour, as measured by the ability to share, was in place by three years of age.

In the pilot study, female children differed across the three trial types. Females were more likely to share, followed by gratifying the self, followed by gratifying others. No difference in performance across the three trial types was found for males. The reasons behind these gender differences remain unclear. It may be hypothesized however, that the gender of the experimenter may have influenced children’s
performance on the task. That is, female children may have been eager to demonstrate their willingness to share with a female experimenter during the shared gratification trials. Hay, Castle, Davies, Demetriou and Stimson (1999) also found gender differences in the prosocial behaviour of 18 to 30 month-old children. The genders did not differ in terms of overall rates of sharing, however, over time girls were much more likely to share with other girls, whereas boys were equally likely to share with boys and girls. Girls in the current study may have also been more competitive with the female experimenter during the other gratification trials. The experimenter noted that female children were often unwilling to allocate stickers to the experimenter if it meant that the number of stickers in the experimenter’s book would be higher than the number of stickers in their own book. Further research with experimenters of both genders could be conducted to confirm this suggestion. However, this pattern was not replicated in the main study.

One issue that must be addressed, however, is that the binary-relational items utilized in the current study did not entail a delay variable. While they are undoubtedly less complex than the ternary-relational items, they do not assess children understanding of all the individual elements that were incorporated in the ternary-relational items. That is, it remains unclear whether the preference by the younger children to choose the immediate rewards was due to their difficulty integrating the three variables required by these ternary-relational items, or potentially a difficulty projecting themselves into the future. Further research is required in order to determine whether the bias of young children against choosing the delay option can be overcome if the complexity of the items involving a delay is reduced. It remains to be seen, however, how this can be achieved without altering the inherent principles underlying the delay of gratification paradigm.
In summary, both the pilot and main study indicated that 3-, 4-, 5- and 6-year-olds passed the binary-relational items, whereas only 5- and 6-year-olds passed the ternary-relational items. These results indicate that success on future-oriented decision-making depends on a child’s ability to process the complex relations in the task.
CHAPTER 6. CONDITIONAL DISCRIMINATION / REVERSAL LEARNING

Research investigating the role of the OFC has often utilized object reversal paradigms. In object reversal learning, the participants learn a simple discrimination between two objects. The choice of one object is rewarded but choice of the other is not. After this discrimination is learned to a criterion, the discrimination is reversed, such that the previously unrewarded object is now rewarded and vice versa. The OFC is thought to be involved in this process through stimulus-reinforcement association learning (Rolls, 2004). That is, it receives information from a variety of sources including taste, somatosensory, visual and olfactory senses, and these are then associated with the primary reinforcers. The OFC also has neurons that detect non-reward, which may be used in reversal learning by resetting the association of the neurons involved (Rolls, 2004). This ability is essential to adaptive functioning, as humans must continually make decisions about which reinforcers to work for, or avoid, depending on their current need states and what rewards have been recently received. Damage to the OFC may consequently impair the correction of behavioural responses when they are no longer appropriate (Rolls, 2000).

Relation of Object Reversal to the OFC

Object reversal is considered to be a hot executive function task because it requires the participant to revise their assessment of whether an object should be approached or avoided (Zelazo & Müller, 2002). As previously stated, the reinforcement value of stimuli is considered to be a primary function of the OFC (Rolls, 2004). Additionally, performance on this task has been found to depend on an intact OFC in non-human animals (Iversen & Mishkin, 1970). OFC-lesioned rhesus monkeys tend to perseverate on the initial discrimination and fail to reverse their responses. Rolls, Hornoak, Wade and McGrath (1994) also found that patients with damage to the OFC perseverated on tasks of reversal learning. Despite often being able to report verbally
that the contingencies had changed, the OFC-lesioned patients were unable to alter their behaviour and continued to respond to the previously rewarded stimulus.

**Development of Object Reversal and Conditional Discrimination Abilities**

Overman, Bachevalier, Schuhmann, and Ryan (1996) examined development in the performance of young children (15 to 54 months) on an object reversal task. On each trial, children were presented with the same two objects and rewarded for reaching to one of them. After a certain number of trials, the contingencies were reversed and children were rewarded for reaching to the other object. The primary dependent variable was the number of trials needed to learn the reversal. Overman et al. (1996) demonstrated age-related improvements in performance on this task in infants and young children. Only the younger female children (less than 30 months) made errors that reflected abnormal perseveration and excessive emotionality.

The presence of gender differences is also consistent with three general findings in the non-human literature. First, 75 day-old male rhesus monkeys have been found to be superior to age-matched females on object-reversal learning (Goldman, Crawford, Stokes, Galkin & Rosvold, 1974). Second, while normal infant males out-perform infant females on the object reversal task, perinatally androgenised infant females perform as well as normal males and better than normal females (Clark & Goldman-Rakic, 1989). Third, early ablations of orbital prefrontal cortex impair object reversal performance of normal infant males and androgenised infant females but such lesions do not impair the performance of normal infant females until much later in life (Clark & Goldman-Rakic, 1989; Goldman, 1971). Overman (2004) stated that such gender differences may be due to a perinatal testosterone surge that may accelerate the functional maturation of OFC in males.

A more complex example of this reversal-learning paradigm is conditional discrimination. This concept was first studied with non-human animals by Lashley
Development of Executive Functions

(1938, cited in Halford, 1982). The basis of this concept is that a choice between two stimuli is made conditional on a further cue, such as the background on which they are displayed. Thus, a triangle might be correct and a square incorrect when the background is white, but when the background is striped, the contingencies are reversed such that square would be correct and the triangle would be incorrect. Transfer can be tested using isomorphic problems with the same structure but different elements (e.g., with shapes such as a cross and a circle and background colours such as blue and green).

While reversal learning tasks involve the initial learning and reversal phases being dealt with in succession, conditional discrimination involves the simultaneous combination of the both learning and reversal phases.

Rudy, Keith and Georgen (1993, Study 2) found that young children were able to master a three-dimensional version of the conditional discrimination task. In this task, the stimuli consisted of two different coloured boxes (A and B), and the correct choice between them depended on the context in which they were presented (either on a plain white partition, or a partition covered with children’s stickers). Three phases were presented: (1) an initial learning phase (A is rewarded, B is not rewarded); (2) a reversal phase (B is rewarded, A is not rewarded); and (3) a conditional discrimination phase (A is rewarded in context 1, B is rewarded in context 2). The results indicated that children younger than 4.5 years of age had no difficulty solving the initial and reversal learning problems, but were unable to reach criterion on the conditional discrimination phase. In contrast, older children were able to reach criterion on all three of the phrases. Thus, the ability to perform conditional discrimination of three-dimensional coloured boxes appeared to develop after 4.5 years. Rudy et al. (1993) attributed this development to the maturation of the hippocampus, however, this might need to be revised in light of the more recent evidence for the involvement of the OFC in processing the reward values of stimuli (Rolls, 2004).
Complexity Analysis of Conditional Discrimination / Reversal Learning

According to the relational complexity approach, conditional discrimination entails a ternary relation because there are three interacting variables (background, object, outcome). The task cannot be performed using processing based on simple associations because each stimulus is associated with both positive and negative outcomes, conditional upon the other stimuli in the set (Halford, 1993). The task can be learned by configural association, however, such processing would not support the transfer of this learning to the subsequent problems. Thus, a transfer problem will initially be included in the design to ensure that relational processing is being used to process the conditional discrimination items.

The object reversal task format described above was used in the current study to model binary-relational items of the conditional discrimination task. The binary-relational reversal learning items will be closely matched to the ternary-relational conditional discrimination items in terms of procedure and format. In each reversal-learning trial, the outcome or reward will be dependent on the object (triangle, circle) while the background variable is held constant. Once the child learns which object is rewarded and which is not, the reinforcement values will be reversed without changing the background variable. These items therefore involve only two variables (object, outcome). The ability to transfer this learning to a subsequent problem that utilises a different stimulus set will also be tested. As aforementioned, transfer to isomorphs is a property of relations, therefore including a transfer phase in the conditional discrimination and object reversal ensures that relational processing, not processing based on associations, is being assessed. Transfer to isomorphs arguably depends on a representation of two mutually exclusive choices, so if it is A it is not B and vice versa. This enables perfect performance on transfer to new problems after one information trial (i.e., one reinforced choice that tells participants which alternative is correct).
Consistent with the relational complexity approach, it was anticipated that 3-, 4-, 5- and 6-year old children would succeed on the binary-relational reversal-learning items. It was also anticipated that only 5- and 6-year-old children would succeed on the ternary-relational conditional discrimination task, while 3- and 4-year-olds would experience difficulty.

Method

Conditional Discrimination Materials and Procedure:

The conditional discrimination task was presented via a Dell latitude laptop computer (30 × 25 cm screen). An example stimulus slide can be seen in Figure 6.1. The two objects (each with a diameter or length of 4cm) were presented side by side on the screen. The objects were white with a black outline. The stimuli (objects and background colours) and outcomes for the two conditional discrimination problems are presented in Table 6.1.

![Example stimulus slide in the conditional discrimination task.](image)

Figure 6.1. Example stimulus slide in the conditional discrimination task.
Table 6.1

*Conditional Discrimination Stimuli and Reward Contingencies*

<table>
<thead>
<tr>
<th>Problem</th>
<th>Background</th>
<th>Object</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1</td>
<td>Red</td>
<td>Triangle</td>
<td>Happy face</td>
</tr>
<tr>
<td></td>
<td>Red</td>
<td>Circle</td>
<td>Sad face</td>
</tr>
<tr>
<td></td>
<td>Blue</td>
<td>Triangle</td>
<td>Sad face</td>
</tr>
<tr>
<td></td>
<td>Blue</td>
<td>Circle</td>
<td>Happy face</td>
</tr>
<tr>
<td>Problem 2</td>
<td>Green</td>
<td>Rectangle</td>
<td>Happy face</td>
</tr>
<tr>
<td></td>
<td>Green</td>
<td>Ellipse</td>
<td>Sad face</td>
</tr>
<tr>
<td></td>
<td>Yellow</td>
<td>Rectangle</td>
<td>Sad Face</td>
</tr>
<tr>
<td></td>
<td>Yellow</td>
<td>Ellipse</td>
<td>Happy face</td>
</tr>
</tbody>
</table>

On each trial, two objects were presented on the computer screen, with varying background colours. The left-right position of the objects was counterbalanced across the stimulus slides. Within each problem, the two objects (e.g., triangle, circle) were presented on one colour background (e.g., red) for 50% of the trials and the other colour background (e.g., blue) for the remaining trials. Children were required to select the object that they thought would produce a happy face. They responded by pressing the right shift key to select the object on the right, or the left shift key to select the object on the left. Feedback was provided after each selection. If the child chose correctly, a happy face appeared on the screen. If they chose incorrectly, a sad face appeared. Children were told at the beginning of the task that they should take notice of the background colour and the shape of the objects and try to choose the object that makes a happy face appear on the screen.

Two conditional discrimination problems were used. Problem 1 was preceded by two ‘information trials’ in which the experimenter explained the feedback to the
children. For example, if the child chose the triangle when the background is red, the experimenter said, “That’s right, the screen was red and you chose the triangle, so the happy face came onto the screen. Try to get another happy face next time.” If the child chose the circle when the background is red, the experimenter said, “That’s not right, the screen was red and you chose the circle, so the sad face came onto the screen. Which one should you have chosen to get a happy face? Try to get a happy face next time, OK?” The displays presented and the feedback received on these trials provided all the relevant information about the background colours, shapes, and their combinations for the current problem.

In the pilot study, two blocks of 12 trials were presented (problem 1). Additional trial blocks were presented if children did not reach the learning criterion of 4 or more correct responses out of 6 for each background colour (based on the binomial distribution using chance = 0.5, this corresponds to \( p = .11 \)). A maximum of six trial blocks were administered. The number of correct responses per block, and number of trial blocks needed to reach criterion on each problem were the primary dependent variables. Problem 2 was then administered using identical task procedures, except that the experimenter did not explain the feedback. The experimenter instead provided non-contingent encouragement aimed at maintaining children’s motivation and interest in the task. Two information trials were presented, followed by 2 blocks of 12 trials, which were repeated if the learning criterion was not met.

In the main study, the minimum number of trials blocks administered in Problem 1 was increased to four. Problem 2 was not presented. See Table 6.2 for a comparison of the trial blocks presented in the pilot and main study.
Table 6.2

Number of Trial Blocks Administered in the Pilot and Main Study of the Conditional Discrimination Task

<table>
<thead>
<tr>
<th>Problem 1</th>
<th>Problem 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>Max.</td>
</tr>
<tr>
<td>Pilot Study</td>
<td>2</td>
</tr>
<tr>
<td>Main study</td>
<td>4</td>
</tr>
</tbody>
</table>

Reversal Learning Stimuli and Procedure

The reversal learning problems were also presented via a Dell laptop computer. Each stimulus slide presented the same objects as they were positioned in the conditional discrimination task, however the background colour remained grey throughout all trials. Table 6.3 details the stimulus and outcomes for the reversal-learning problems.

Table 6.3

Reversal Learning Stimuli and Reward Contingencies

<table>
<thead>
<tr>
<th>Phase</th>
<th>Background</th>
<th>Shape</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1:  Learning</td>
<td>Grey</td>
<td>Triangle</td>
<td>Happy face</td>
</tr>
<tr>
<td></td>
<td>Grey</td>
<td>Circle</td>
<td>Sad face</td>
</tr>
<tr>
<td>Reversal</td>
<td>Grey</td>
<td>Triangle</td>
<td>Sad face</td>
</tr>
<tr>
<td></td>
<td>Grey</td>
<td>Circle</td>
<td>Happy face</td>
</tr>
<tr>
<td>Problem 2:  Learning</td>
<td>Grey</td>
<td>Rectangle</td>
<td>Happy face</td>
</tr>
<tr>
<td></td>
<td>Grey</td>
<td>Ellipse</td>
<td>Sad face</td>
</tr>
<tr>
<td>Reversal</td>
<td>Grey</td>
<td>Rectangle</td>
<td>Sad face</td>
</tr>
<tr>
<td></td>
<td>Grey</td>
<td>Ellipse</td>
<td>Happy face</td>
</tr>
</tbody>
</table>
The reversal learning items were presented in a manner similar to the conditional discrimination items. Two objects were presented on the screen with a constant background colour. Children were instructed to press the left or right shift key to indicate their choice of object. Feedback in the form of happy and sad faces followed correct and incorrect choices respectively. The learning phase consisted of one block of 12 trials. The trial block was repeated for children who failed to reach the learning criterion of 9 out of 12 correct responses ($p = .019$). A maximum of two trial blocks were provided in the learning phase. Once the learning criterion was reached, the reversal phase was presented, in which the reinforcement contingencies were reversed. The reversal phase consisted of 1 block of 12 trials, which was repeated for children who failed to reach the learning criterion of 9 out of 12 correct responses ($p = .019$). Problem 2 of the reversal learning problem was then be presented, during which the experimenter gave no explanation of the feedback. The number of correct responses across the first block of each phase and the number of trials needed to learn the reversal were the primary dependent variables for the binary level items. As can be seen in Table 6.4, Problem 2 was not administered in the main study and the minimum number of trial blocks was increased to 2 for all phases.

All children completed both the reversal learning (binary-relational) and conditional discrimination (ternary-relational) items in a counter-balanced order. The binary-relational and ternary-relational items were closely matched in terms of stimuli, task instructions and procedures.
Table 6.4

Number of Trial Blocks Administered in the Learning and Reversal Phases of Problem 1 and 2 of the Reversal Learning Task

<table>
<thead>
<tr>
<th></th>
<th>Problem 1</th>
<th>Problem 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Learning</td>
<td>Reversal</td>
</tr>
<tr>
<td>Pilot Study</td>
<td>1  2</td>
<td>1  2</td>
</tr>
<tr>
<td>Main study</td>
<td>2  2</td>
<td>2  2</td>
</tr>
</tbody>
</table>

Results

Pilot Study

Conditional Discrimination. Figure 6.2 shows the cumulative number of children in each age group who had reached criterion (4 or more correct out of 6 on each background colour) after each trial block of problems 1 and 2. Whereas 87.50% of 5-year-olds had met the criterion after the first trial block of problem 1, 87.50% of 3-year-olds had not reached the learning criterion by the 6th trial block of problem 1, and 80% had not done so after the 4th trial block of problem 2. Of 4-year-olds, 41.67% reached criterion by trial block 6 of problem 1, and 54.16% of 4-year-olds reached criterion by trial block 4 of problem 2.
Figure 6.2. Cumulative number of children who met criterion on each trial block on conditional discrimination task (percentage of children who met criterion by block 6 of problem 1; $N = 72$).

Reversal Learning. Figure 6.3 shows the cumulative number of children in each age group who had reached criterion (9 or more out of 12) after each trial block in the learning and reversal phases of problems 1 and 2. On problem 1, at least 60% of children in each age group had reached the learning criteria after the first trial block and at least 87.5% had done so after trial block 2 in the learning phase. When the contingencies were reversed, at least 87% in each age group again reached the criterion after two trial blocks. On problem 2 at least 75% of children in each age group reached the learning criteria on the first trial block and 96% had done so after trial block 2 of the learning phase. When the contingencies were reversed, at least 96% in each age group again reached the criterion after two trial blocks.
Figure 6.3. Cumulative number of children who reached the learning criteria on each block of the learning and reversal phases of problem 1 and 2 (N = 72).

Number of items correct. The number of items children performed correctly on the first block of the learning and reversal trials (out of 24) for the reversal-learning problems, and the first two trial blocks of the conditional discrimination items (out of 24) was calculated. The results for problems 1 and 2 was analysed using a 3 (age: 3, 4, 5 years) × 2 (gender: male, female) × 2 (complexity) × 2 (problem) ANOVA. The analysis indicated significant main effects of complexity, $F(1, 66) = 153.19, p < .001, \eta^2 = .699$, age, $F(2, 66) = 56.55, p < .001, \eta^2 = .631$, and problem, $F(1, 66) = 7.06, p = .010, \eta^2 = .097$. Children performed better on problem 2 ($M = 18.33, SE = .32$) than they did on problem 1 ($M = 17.42, SE = .29$). No main effect or interaction involving gender was recorded.

The main effects of age and complexity were modified by a significant Complexity × Age interaction, $F(2, 66) = 22.62, p < .001, \eta^2 = .407$. Means and standard errors for the interaction are presented in Figure 6.4. Analysis of simple effects indicated that the performance increased significantly across the three age groups on both the reversal learning, $F(2, 69) = 12.79, p < .001, \eta^2 = .271$, and conditional discrimination items, $F(2, 69) = 70.51, p < .001, \eta^2 = .671$. The magnitude of this effect
however, was much stronger on the ternary-relational items than the binary-relational items. For the reversal learning items, a Scheffe test indicated significant differences between the means for: 3-year-olds and 5-year-olds \((p < .001)\); 4-year-olds and 5-year-olds \((p < .001)\). For the conditional discrimination items, a Scheffe test indicated significant differences between the means for 3-year-olds and 4-year-olds \((p = .005)\), and 3-year-olds and 5-year-olds \((p < .001)\); 4-year-olds and 5-year-olds \((p < .001)\).

![Mean number of correct choices on the reversal learning (trial block 1 of learning trials and trial block 1 of reversal trials) and conditional discrimination items (trial blocks 1 and 2), problems 1 and 2 combined \((N = 72)\).](image)

**Figure 6.4.** Mean number of correct choices on the reversal learning (trial block 1 of learning trials and trial block 1 of reversal trials) and conditional discrimination items (trial blocks 1 and 2), problems 1 and 2 combined \((N = 72)\). Single sample \(t\)-tests were used to compare the means for each age group to their respective chance levels \((12)\) for the reversal learning and conditional discrimination scores. For reversal learning, the mean scores of 3-, 4-, and 5-year-olds were all significantly above chance on Problem 1, smallest \(t(23) = 8.85, p < .001\), and Problem 2, smallest \(t(23) = 11.98, p < .001\). For the conditional discrimination items, the mean scores on Problem 1 were at chance level only for the 3-year-olds and 4-year-olds, however were significantly above chance for the 5-year-olds, \(t(23) = 12.74, p < .001\). For Problem 2, the mean scores were at chance level only for the 3-year-olds, and
significantly above chance for the 4-year-olds, $t(23) = 3.10, p = .005$ and 5-year-olds, $t(23) = 12.74, p < .001$.

Performances of the individual children were evaluated against chance. The probability of a correct response on each trial was 0.50, thus scores of 16 or more out of 24 are significantly above chance according to the binomial table ($p < .05$). As can be seen in Table 6.5, the majority of children from all age groups performed above chance levels on both problems 1 and 2 of the reversal learning task. A minority of 3- and 4-year-olds and a majority of 5-year-olds performed significantly above chance on the both Problems 1 and 2 of the conditional discrimination task. The pass-fail data indicated that the median age of attainment was three years (or earlier) for the reversal learning items, but not until five years for the conditional discrimination items.

Table 6.5

*Frequency of Children Within Each Age-Group Performing above Chance Levels on the Binary- and Ternary Relational Items of the Learning Phase ($N = 72$).*

<table>
<thead>
<tr>
<th>Age</th>
<th>Reversal Learning</th>
<th></th>
<th>Conditional Discrimination</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Problem 1</td>
<td>Problem 2</td>
<td></td>
<td>Problem 1</td>
</tr>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>3</td>
<td>21</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>4-years</td>
<td>3</td>
<td>21</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>5-years</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 6.6 relates the pass-fail scores for problems 1 and 2 of the reversal learning and conditional discrimination items. For problem 1, 51% of children...
performed consistently, failing or passing items at both levels of complexity. The remaining children showed an inconsistent pattern. These children were significantly more likely to pass the reversal learning items and fail the conditional discrimination items (100%), than to show the reverse pattern, McNemar test, $\chi^2(1, N = 72) = 33.03, p < .001$. For problem 2, 54% of children performed consistently, failing or passing items at both levels of complexity. The remaining children showed an inconsistent pattern. These children were significantly more likely to pass the reversal learning items and fail the conditional discrimination items (100%), than to show the reverse pattern, McNemar test, $\chi^2(1, N = 72) = 31.03, p < .001$.

Table 6.6

*Crosstabulation of Children Passing and Failing the Binary- and Ternary-Relational Items of the Conditional Discrimination Task Learning Phase (N = 72).*

<table>
<thead>
<tr>
<th></th>
<th>Problem 1</th>
<th>Problem 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Binary</td>
<td>Ternary</td>
</tr>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>Fail</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Pass</td>
<td>35</td>
<td>30</td>
</tr>
</tbody>
</table>

*Main Study*

*Conditional Discrimination.* Figure 6.5 shows the cumulative number of children in each age group who had reached criterion after each trial block. Whereas 56% 5-year-olds and 60% of 6-years-olds had met criterion by Trial Block 1, 83.33% of 3-year-olds had not reached the learning criterion by the 6th trial block. Of the 4-year-
olds, 20% reached criterion by trial block 1, and 73.33% had reached criterion by trial block 6.

**Figure 6.5.** Cumulative number of children who met criterion on each trial block on conditional discrimination task (percentage of children who met criterion by block 6).

*Reversal Learning.* Figure 6.6 shows the cumulative number of children in each age group who had reached criterion after each trial block. At least 80% of children in each age group had reached the learning criterion after the first trial block and 100% had done so after trial block 2. When the contingencies were reversed, 40% of 3-year-olds reached the criterion after the first trial block, while at least 80% of the other age groups did so. At least 96.67% in each age group again reached the criterion after two trial blocks.
Figure 6.6. Cumulative number of children in each age group who reached the learning criteria on the learning and reversal problems (N = 120).

Number of Items Correct

The number of items performed correctly was based on the first two trial blocks of the learning phase and first two trial blocks of the reversal phase (maximum = 48) for the reversal learning items, and the first four trial blocks (maximum = 48) of the conditional discrimination problem. The distribution of the 3-year-olds on the conditional discrimination was not normally distributed. Inspection of the raw scores indicated that this was due to the performance of two children who performed better than expected on the ternary-relational items. The removal of these scores did not significantly impact the findings. Thus, all data points were retained in the final analysis reported below.

The order in which the conditional discrimination and reversal learning tasks were presented had no effect so this variable was excluded from further analysis, $F(1, 118) = 0.54, p = .464, \eta^2 = .005$. The number of correct items was analysed using a 4 (age: 3, 4, 5, 6 years) $\times$ 2 (gender: male, female) $\times$ 2 (complexity: reversal learning, conditional discrimination) ANOVA.
The ANOVA indicated significant main effects of complexity, $F(1, 112) = 213.68, p < .001, \eta^2 = .656$, and age, $F(3, 112) = 76.06, p < .001, \eta^2 = .671$. A significant Complexity $\times$ Age interaction was found, $F(3, 112) = 19.97, p < .001, \eta^2 = .349$. Means and standard errors for the interaction are presented in Figure 6.7. Analysis of simple effects indicated that the performance of children increased significantly across the four age groups on both the reversal learning, $F(3, 119) = 27.17, p < .001, \eta^2 = .413$, and conditional discrimination items, $F(3, 119) = 55.79, p < .001, \eta^2 = .591$. For the reversal learning items, a Scheffe test indicated significant differences between the means for: 3-year-olds and 4-year-olds ($p = .001$), and 3-year-olds and 5- and 6-year-olds ($p$’s < .001); 4-year-olds and 5-year-olds ($p = .023$) and 4-year-olds and 6-year-olds ($p = .004$). For the conditional discrimination items, a Scheffe test indicated significant differences between the means for: 3-year-olds and 4-year-olds ($p = .005$), and 3-year-olds and 5- and 6-year-olds ($p$’s < .001); 4-year-olds and 5- and 6-year-olds ($p$’s < .001).
Single sample $t$-tests were used to compare the means for each age group to their respective chance levels (24) for the reversal learning (trial blocks 1 and 2 of learning and reversal phases) and conditional discrimination scores (trial blocks 1 to 4). For the reversal learning items, the mean scores of 3-, 4-, 5-, and 6-year-olds were all significantly above chance, smallest $t(29) = 21.33, p < .001$. For the conditional discrimination items, the mean scores were at chance levels for 3-year-olds, $t(29) = 1.07, p = .295$. The mean scores of 4-, 5- and 6-year-olds were significantly higher than chance ($p$’s < .001).

Performance of the individual children was evaluated against chance for the reversal learning (trial blocks 1 and 2 of learning and reversal phases) and conditional discrimination scores (trial blocks 1 to 4). The probability of a correct response on each item was 0.50, thus scores of 30 or more out of 48 were significantly above chance according to the binomial table ($p < .05$). As can be seen in Table 6.7, all children performed above chance levels on the reversal learning items. A minority of 3- and 4-year-olds and a majority of 5- and 6-year-olds performed significantly above chance on the conditional discrimination items. The pass-fail data indicated that the median age of attainment was five years for the conditional discrimination task, and three years (or earlier) for the reversal learning task.

Table 6.8 relates the pass-fail scores for the reversal learning and conditional discrimination items. Sixty-one percent of children (73) performed consistently, failing or passing items at both levels of complexity. The remaining (47) children showed an inconsistent pattern. These children were significantly more likely to pass the reversal learning and fail the conditional discrimination items (100%), than to show the reverse pattern, McNemar test, $\chi^2(1, N = 120) = 45.02, p < .001$. 


Table 6.7

*Frequency of Children Within Each Age-Group Performing above Chance Levels on the Reversal Learning and Conditional Discrimination Items (N = 120).*

<table>
<thead>
<tr>
<th>Age</th>
<th>Reversal Learning</th>
<th>Conditional Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>4-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>5-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>6-years</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 6.8

*Crosstabulation of Children Passing and Failing the Reversal Learning and the Conditional Discrimination Items (N = 120)*

<table>
<thead>
<tr>
<th>Binary</th>
<th>Ternary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
</tr>
<tr>
<td>Fail</td>
<td>0</td>
</tr>
<tr>
<td>Pass</td>
<td>47</td>
</tr>
</tbody>
</table>
Discussion

The results of problem 1 of both the pilot and main study indicate significant development in the performance of object reversal / conditional discrimination tasks during the preschool years. Firstly, the majority of children in all age groups passed the reversal-learning task. This is consistent with the results of Overman et al. (1996) who demonstrated that children as young as 30 months of age succeeded on an object reversal task. In contrast, the majority of 3-year-olds did not pass the conditional discrimination task, while the majority of 5- and 6-year-olds succeeded. The mean score of 4-year-olds was significantly above chance on these items, however, only 38.90% of 4-year-olds succeeded on the conditional discrimination items. This also supports the findings of Rudy et al. (1993) who found that only children greater than 4 years, 9 months of age achieved criterion on a 3-dimensional conditional discrimination task.

These results support the relational complexity interpretation. The binary-relational object reversal items involved only two variables and as such were passed by 3-, 4-, 5- and 6-year-old children. The conditional discrimination items, however, involved three interacting variables. As such, only children five years and older processed the ternary-relations required for the conditional discrimination task. While 3- and 4-year-olds processed the component binary relations involved in these tasks, it appears they are unable to integrate these binary relations into a ternary relation.

The results of the second problem in the pilot study also provide support for the use of relational processing in the performance of conditional discrimination tasks. All age groups passed problem 2 of the reversal learning task, with even 75% of 3-year-olds reaching criterion in the first block of the transfer problem. In contrast, the majority of 3-year-old children did not reach criterion on either problem 1 or problem 2 of the conditional discrimination task. The performance of 4-year-olds improved across the two problems, with 41.67% of children reaching criteria on problem 1, and 54.16% of
4-year-olds reaching criteria on problem 2. Five-year-olds were also able to transfer their learning to the second problem. The performance of 5-year-olds was close to ceiling for both problems. The ability of 4-year-old children to transfer this knowledge in the pilot study to a subsequent problem that utilised a different stimulus set suggested that children were in fact solving this problem through the use of relational processing, rather than processing based on associations.

It must also be noted that both the pilot and main study failed to find any effect of gender on either the reversal learning or conditional discrimination tasks. The equivalent performance of males and females on the reversal learning task appears inconsistent with the work of both Overman et al. (1996) with human infants, and the non-human animal literature (e.g., Goldman et al., 1974; Clark & Goldman-Rakic, 1989). Both of these lines of evidence have suggested a male superiority effect in both young human infants and animals on such learning tasks, which was suggested to be attributable to perinatal testosterone surges that lead to an acceleration of the maturation of the OFC. However, the work of Overman et al (1996) found that the gender difference for human infants was only present for children less than 30 months of age. The older age groups of the children included in the current study (minimum = 36 months) may have precluded such differences being found, suggesting that the earlier male superiority may have leveled out by three years of age.

In summary, both the pilot and main study indicated that 3-, 4-, 5- and 6-year-olds passed the reversal learning task, whereas only 5- and 6-year-olds passed the conditional discrimination task. These results indicate that the primary requirement for successful completion of the conditional discrimination items is children’s ability to process complex relations.
An additional domain of hot executive function that has been studied extensively in the literature is that of ‘theory-of-mind’. This term refers to a set of explicit and interconnected concepts for representing mental states (Astington, Harris & Olson, 1988). Research into theory-of-mind development has largely focused on the preschool years, when children begin to recognise mental states, including beliefs, desires, intentions and emotions, in themselves and others. The development of a theory-of-mind provides a corner stone for social and intellectual behaviour by allowing the child to understand social interactions through the attribution of mental states to themselves and others (Astington et al., 1988).

*Relation of Theory-of-Mind to the OFC*

Theory-of-mind is considered a hot executive domain because it requires flexible problem representation and reversal of responding in the context of concrete social information (Zelazo & Müller, 2002). The regulation of social behaviour, broadly defined, has consistently been implicated as a function of the OFC (Damasio, 1994). The specific involvement of the OFC in theory-of-mind abilities, however, remains somewhat speculative. A study by Stone, Baron-Cohen and Knight (1998) compared the performance of patients with bilateral damage to OFC ($N = 5$) and unilateral damage to the left DL-PFC ($N = 5$) on a series of theory-of-mind tasks varying in difficulty. Patients with bilateral damage to the OFC were impaired on a sophisticated test of theory-of-mind that involved recognition of faux pas. In contrast, the dorsolateral lesioned patients had no specific theory-of-mind impairment and experienced difficulty only on the versions of the tasks requiring significant working memory resources. The authors concluded that the OFC is part of a distributed circuit involved in theory-of-mind and that it may be particularly involved in tasks with an affective component.
Recent neuroimaging studies, however, have also implicated a role of the medial prefrontal cortex in inferencing. Goel, Grafman, Sadato and Hallet (1995) used PET to scan adults as they made inferences about objects that required either a visual description of the object, memory retrieval, an inference from the object’s form to its function, or an inference that required modeling another person’s mental state. The results indicated widespread activation of the left medial frontal lobe and left temporal lobe during the task requiring mentalistic inferences. Such conflicting findings have led researchers to propose a combined mediofrontal-orbitofrontal circuitry responsible for the control of theory-of-mind functioning. For example, Sabbagh (2004) hypothesized that theory-of-mind reasoning can be fractionated into two distinct neural circuits. While the orbital cortex may be especially important for decoding mental states, the left medial frontal regions may be important for reasoning about mental states. Based on an fMRI study with healthy participants, Hynes, Baird and Grafton (2006) argued for a multiple component model of theory-of-mind functioning. Group results demonstrated that the medial orbitofrontal lobe was preferentially involved in an emotional perspective-taking task, while the cognitive perspective-taking condition engaged more lateral and anterior regions of the ventral frontal lobes. Thus, the OFC may be part of a component system involved in the processing of social stimuli, including theory-of-mind abilities and other hot executive functions.

Amodio and Frith (2006) further reviewed studies that had examined three broad categories of tasks thought to involve the medial frontal cortex: a) the control and monitoring of action, b) the monitoring of outcomes that relate to punishments and reward, and c) social cognition such as self-reflection, person perception, and making inferences about others’ thoughts. On the basis of this evidence, the authors proposed a theoretical model of medial frontal cortical function that incorporated different aspects of social cognitive processing. Overall, the more caudal region of the medial frontal
cortex was associated with actions, whereas the orbital region was associated with outcomes. The anterior region of the medial frontal cortex was linked to higher cognitive processes that enable reflection on the values linked to outcomes and actions. The different functions across these regions are suggested to form a systematic map that integrates the different aspects of social cognitive processing.

Assessment of Theory-of-Mind Development

Research investigating children’s acquisition of a theory-of-mind has largely employed three types of assessment tasks: a) false-belief; b) representational change; and c) appearance-reality.

Wimmer and Perner (1983) devised a set of false belief tasks to assess the acquisition of a theory-of-mind in young children. This paradigm explored children’s understanding that people will act in accordance with their own beliefs, even if those beliefs are false. For example, in their change-of-location tasks, children are presented with a story in which a character places an object in location A. The character is removed from the scene, and during his absence the object is relocated to location B. The child is then asked to report where the character will look for the object upon their return. To pass this task, the child must acknowledge that the character will act on the basis of their own belief about the object’s location, rather than the object’s actual location.

Recently, a meta-analysis was conducted on 178 separate studies related to false-belief understanding that were published prior to January 1998. It addressed the empirical inconsistencies regarding the explanation for, and the age of emergence of, false-belief understanding (Wellman et al., 2001). The analysis indicated that false-belief performance showed a consistent developmental pattern, even across various countries and various task manipulations. Before 41 months (3 years, 5 months) children performed below chance, making the classic false-belief error, that is, reporting
that the character will look in the location to which the object has been moved. After 48 months (4 years), the children performed above chance levels. Average performance changed rapidly during the period 3 to $4\frac{1}{2}$ years from consistently wrong to consistently correct.

Representational change tasks assess children’s understanding of their own false-belief. In these items, children are not asked about a character’s false beliefs, but rather about their own (Wellman et al., 2001). For example, a child may be shown a closed crayon box (assumed to contain crayons) that is then opened to reveal that it actually contains candies. The child is then asked, “Before you looked inside, did you think the box contained crayons or candies?” To pass these items, the child must be able to differentiate between their earlier false belief, and their new knowledge of the box’s true contents, by acknowledging that they initially believed that the box contained crayons. Gopnik and Astington (1988) demonstrated that 3-year-olds did not recognise that their own beliefs changed after they found out that they were initially wrong. This was evident in their tendency to report that they had known the newly learned information for a long time. Thus, the 3-year-old children claimed to have always known the crayon box contained candies, even though they only acquired this knowledge after looking inside the box. It was not until approximately four years of age that the children began to appreciate their own representational change (Gopnik & Astington, 1988). Wellman et al.’s (2001) meta-analysis of 235 conditions, also found that children’s correct responses to false-belief questions for self versus other did not differ, and were virtually identical at the younger ages.

A range of appearance-reality tasks investigated by Flavell, Flavell and Green (1983) also examined children’s ability to make inferences about their own and others’ mental states. In these tasks, children are asked to distinguish between how things appear and how or what they really are. For example, a child is shown an object viewed
directly and then when it has been covered by a coloured blue filter. The child is then asked what colour the object “looks like” and what colour it is “really and truly”. To pass, children must be able to appreciate that an object could simultaneously appear one way (blue) while in reality be another way (white). Flavell et al. (1983) demonstrated that on a series of appearance-reality items, the 4-year-olds were much better at making this distinction than 3-year-olds, and that the 5-year-olds performed almost errorlessly. The authors also reported that 3-year-olds demonstrated conceptual difficulties with the appearance-reality distinction in two different ways: (1) by incorrectly reporting appearance when asked to report reality (labeled “phenomenism”); or (2) by incorrectly reporting reality when asked to report appearance (“intellectual realism”). Thus the emergence of false-belief, representational-change and appearance-reality understanding at around 4- to 5-years-of-age is a much replicated and robust finding (Perner, Leekam, & Wimmer, 1987; Gopnik & Astington, 1988; Flavell, Green, & Flavell, 1990)

It has been suggested that the reason that 3-year-olds fail on theory-of-mind tasks may be due to a number of methodological issues, such as the level of verbal ability or memory demands required by the experimental procedures. However, three lines of evidence have emerged against methodological explanations, namely: a) attempts to improve the performance of 3-year-olds by reducing tasks demands have largely failed (e.g., Call & Tomasello, 1999; Perner et al., 1987; Wimmer & Hartl, 1991); b) 3-year-olds from different cultures perform similarly poorly on the same tasks (Avis & Harris, 1991; Flavell, Zhang, Zou, Dong & Qi, 1983); and c) attempts to train children on these tasks have also proven unsuccessful (Flavell, Green, & Flavell, 1986). In comparison, the meta-analysis Wellman et al. (2001) concluded that several task manipulations might increase young children’s performance on false-belief tasks. Specifically, framing the task in terms of explicit deception or trickery, involving the
child in actively making the key transformations, and highlighting the salience of the character’s mental state may all help improve the performance of young children to some extent (Wellman et al., 2001). In general, however, the difficulty experienced by 3-year-old children on false-belief and appearance-reality tasks is a genuine and deep-seated conceptual one.

Relation of Theory-of-Mind to Executive Functions

Recent research has demonstrated significant correlations between cool executive function skills and performance on theory-of-mind tasks. Hughes (1998) investigated the proposed relationship between children’s executive function and their developing theory-of-mind capabilities. In this study, three aspects of executive function were first distinguished: working memory (as measured by a visual search task and an auditory sequencing task), attentional flexibility (assessed by a colour/shape shifting task and a magnets pattern-making task) and inhibitory control (performance on ‘detour-reaching box’ task and a ‘fist and finger’ hand game). Specific links were found between these three components and theory-of-mind performance in 3- and 4-year-olds, in particular between working memory and false belief-prediction, and between inhibitory control and both deceit and false-belief explanation. In a study that was described in Chapter 2, Carlson and Moses (2001) found that theory-of-mind in 3- and 4-year-olds was significantly correlated with executive function measures, especially those that involved conflict. A range of other studies have also reported similar correlations between these two domains (e.g., Hughes, Dunn & White, 1998; Perner & Lang, 2000).

The development of theory-of-mind and executive function abilities in young children have been hypothesised to be related in five possible ways (Perner & Lang, 1999). Firstly, as described above, theory-of-mind and executive functions may be mediated by similar regions of the prefrontal cortex. Secondly, an intact theory-of-mind
Development of Executive Functions

may be a prerequisite for executive function skills. That is, the development of theory-of-mind skills allows children a better understanding of their own mentality which leads to greater control of their mental processing abilities. Thirdly, there may be executive components in theory-of-mind tasks. For example, typical theory-of-mind tasks may require the suppression of natural response tendencies, such as pointing to the true location of the hidden objects in change-of-location false-belief tasks (Perner & Lang, 1999). Fourthly, emergence accounts claim that a certain level of executive ability needs to be in place before a belief concept can even be constructed (Moses, 2001). To this end, Carlson, Mandell and Williams (2004) conducted a longitudinal study that examined the relations between executive functions and theory-of-mind task performance in 81 children at 24 and 39 months of age. Executive function and theory-of-mind tasks were significantly related at 39 months of age however, no significant relation was evident at Time 1. Additionally, executive function scores are Time 1 significantly predicted theory of mind performance at Time 2, independent of several controls. The authors concluded that executive functions appear to be an important, although not exclusive, contributor to children’s developing understanding of mind.

The final proposed link between theory-of-mind and executive function skills suggest that age-related improvements in both domains may be related because each reflects maturation of an underlying domain-general developmental process, such as the level of complexity that the child is able to process (Frye et al., 1995). The argument encompassed by this domain-general view is that advances in children’s understanding of the mind show important relations with other aspects of cognitive development. For example, Gordon and Olson (1998), attempted to account for children’s acquisition of a theory-of-mind in terms of a series of incremental changes in their capacity for “holding in mind”. Their results demonstrated a close relation between children’s performance on standard theory-of-mind tasks and changes in computational resources as measured by
two dual tasks (a counting and labeling task, and a finger tapping and labeling task). While controlling for the effects of age, the counting and labeling task accounted for 22% of variance in the false-belief and 18% of the variance in appearance-reality, while the finger tapping and labeling task accounted for 14% of the variance in false-belief and 19% of the variance in appearance-reality.

Davis and Pratt (1995) also found a significant relationship between working memory capacity (indicated by a backward digit span task) and false belief performance, even when age and language ability were statistically controlled. The authors concluded that children may have an innate predisposition for mentalistic thinking, but this ability is limited at any given point in development by the child’s working memory capacity. These findings were strengthened by the work of Keenan, Olson and Marini (1998) which employed a wider range of false-belief tasks as well as a backward digit span task. The results indicated that working memory was a significant predictor of children’s performance on the false-belief tasks, uniquely accounting for 7.4% of the variance. Thus, while the findings indicate a role for working memory in the development of false-belief understanding, it is only one of a number of factors that underlie the development of children’s theory-of-mind (Keenan et al., 1998).

Complexity Explanations of Theory-of-Mind Development

As described in Chapter 2, two recent approaches have applied complexity interpretations to the development of theory-of-mind. According to the CCC theory, theory-of-mind tasks require young children to reason based on embedded rules, as they need to consider not only their own mental states and perspectives, but also those of another person. Use of embedded rules make it possible to switch between judgments based on these two conflicting perspectives. Frye et al. (1995) found that advances in theory-of-mind (appearance-reality, false-belief and representational change items), the DCCS, and a physical causality task between the ages of three and five depended on
being able to switch judgments across conditions, and that reasoning by embedded rules could account for these changes. In Experiment 3, after the effects of age were partialled out, the DCCS accounted for 10% of variance in false-belief and 6% of the variance in representational change.

Müller, Zelazo and Imrisek (2005) also found that performance on the DCCS predicted performance on false-belief tasks in 69 3- to 5-year-old children, even after controlling for children’s age and verbal ability. The authors concluded that the developmental relation between theory-of-mind and executive function exists as both task types depend on a common processing element, representational flexibility. That is, both sets of tasks require either shifting from one sorting dimension to another (as in the DCCS) or in shifting between cognitive perspectives (as in theory-of-mind tasks). Thus, the ability to switch judgments across conflicting conditions is viewed as an essential cognitive capacity underlying performance of both executive function and theory-of-mind task.

Using relational complexity theory, Halford (1993) argued that variations in performance on theory-of-mind tasks reflect in part the complexity of the mental representations that these tasks require. To succeed on appearance-reality tasks, children must appreciate that the relation between an object, and a person’s perception of that object is conditional upon a third variable, viewing condition (see Figure 7.1). Thus, the appearance-reality tasks are equivalent to a ternary relation, which Halford et al. (1998) expressed as: Seen-object (<condition>,<object-attribute>,<object-percept>).

Using the previous appearance-reality example in which a white object was covered by a blue filter, the relevant relational instances can be expressed as:

a) Seen-object (no filter, object-white, percept-white), and

b) Seen-object (blue-filter, object-white, percept-blue)

where a) and b) corresponds to reality and appearance respectively.
Figure 7.1. Relational complexity analysis of appearance-reality, false-belief, connections and transformation tasks.

Thus, while 3-year-olds should be able to process the component binary relations involved in these tasks, they should experience difficulty integrating these binary relations into a ternary relation. Andrews et al. (Exp. 1 & 2, 2003) examined the performance of normally developing children on standard theory-of-mind tasks (appearance reality and false belief items) as well as on some tasks from the cool content domains (transitivity, cardinality, and class inclusion classification) that involve binary and ternary items. Significant correlations were found between the theory-of-mind items and other ternary-relational tasks from the cool domain. Performance on relational complexity tasks from cool domains accounted for 85% of the age-related
variance on theory-of-mind tasks, and an additional 11.6% of the variance independent of age.

Further, in Experiment 3 Andrews et al. (2003) also introduced a complexity manipulation into the theory-of-mind tasks. The study examined the performance of 3-, 4- and 5-year-olds on the same theory-of-mind tasks as well as connections and transformation tasks, which involve similar content to theory-of-mind tasks but have a lower level of complexity. The connections tasks examine whether children can specify whether objects in the environment are perceptible to themselves or to others. As can been seen in Figure 7.1, this involves a link between two variables, the environmental cue and the person’s representation. Transformation tasks examine children’s understanding of the transformation applied in theory-of-mind tasks, and involve the link between the setting condition and the person’s representation. Thus, both connections and transformation tasks were binary-relational. Significant complexity effects were found for the theory-of-mind tasks. While the binary-relational items were passed by 3-, 4- and 5-year-old children, the ternary-relational items were not passed until a median age of 5-years. A predictor task from a different domain (class inclusion) which contained items at two levels of complexity (binary-relational, ternary-relational) was also included. The predictor task accounted for more than 80% of age-related variance in theory-of-mind.

The main study administered a battery of theory-of-mind tasks presented at two levels of complexity (binary-relational and ternary-relational). Consistent with the predictions in Chapter 3, the majority of children in all age-groups were anticipated to pass the binary-relational connections and transformation items. In contrast, only 5- and 6-year-olds were anticipated to perform above chance levels on the ternary-relational false-belief and appearance-reality items.
Method

The tasks, materials and procedures utilized in the main study were based on Andrews et al. (Experiment 3, 2003). The order of the tasks, individual items and subsequent questions were presented in a counter balanced order. Responses for the connections and transformation items, and appearance-reality and false-belief items were recorded by the experimenter on response sheets (see Appendix D and E respectively).

Connections and Transformation Tasks

Two connections items and four transformation items were presented.

Connections: Drum-barrier (Flavell et al., 1990). The children were shown a drum and told, “Here is a drum”. The drum was sounded and then placed behind a white cardboard barrier positioned such that the experimenter could see the drum but the child could not. The test questions (correct answers) were: (1) “Look, do you see the drum right now?” (no); and (2) “How about me. Do I see the drum right now?” (yes). The position of the drum was changed so that the child could see the drum, but the experimenter could not. (1) “Look, do you see the drum right now?” (yes); and (2) “How about me. Do I see the drum right now?” (no).
Connections: Open-Ended Box. The children were shown a crayon box and told, “Here is a crayon box”. The crayon box was positioned so that the open end faced toward the child, and they were asked: (1) “Look, can you see what is inside the box right now?” (yes); and (2) “How about me? Can I see what is inside the box right now?” (no). The box was then positioned so that the open end faced towards the experimenter, and the child was asked, (1) “Look, can you see what is inside the box right now?” (no); and (2) “How about me. Can I see what is inside the box right now?” (yes).

Each connections item was scored as correct if responses to the component questions (1 and 2) within each display orientation were both correct (maximum = 4).

Transformation: Colour. A yellow plastic filter was held up for the children to see through, and they were asked, “What colour do things look through this plastic, yellow or red?” (yellow). A purple filter was then held up for children to look through, and they were asked, “What colour do things look through this plastic, green or purple?” (purple).

Transformation: Wobbly. A container of water was held up for the children to see through, and they were asked, “How do things look through this water, still or wobbly?” (wobbly). The children were then given a clear glass and asked, “How do things look through this glass, still or wobbly?” (still).

Transformation: Change in contents. A closed biscuit box was placed in the children’s view, and they were asked, “Let’s pretend that someone took everything out of this box, and then put some paper in this box. Can you pretend that happened?” (1) “What would be in the box if that happened, biscuits or paper?” (paper). “Let’s pretend that someone took everything out of this box and then put some stones in this box. Can you pretend that happened?” (2) “What would be in the box if that happened, stones or biscuits?” (stones).
Development of Executive Functions 143

**Transformation: Change in location.** Three boxes, representing locations A, B, and C were placed in front of the children. A toy dog was placed at location A, and the children were told, "Here is a toy dog. Let’s put it here (location A). Let’s pretend that someone picked up this toy dog and put it there (point to location B). Can you pretend that happened?" (1) "Where would the dog be if that happened, A or B?" (B). The children were then shown a toy bear at location A and told, "Here is a toy bear. Let’s pretend that someone picked up this bear and put it over there (point to location C). Can you pretend that happened?" (2) "Where would the bear be if that happened, A or C?" (C).

Each transformations item was scored as correct if responses to the component questions were both correct (maximum = 4). Thus, the maximum score for the binary-relational theory-of-mind items (connections and transformation combined) was 8.

**Appearance-Reality and False Belief Tasks**

Three appearance-reality items and three false-belief items were used:

*Fish-filter task* (Flavell et al., 1983). Children were shown a white cardboard fish covered by a blue filter so that the fish appeared blue. The filter was removed revealing the fish’s real colour to be white. The children were then asked to report the colour of the fish and given feedback on their responses. The filter was replaced over the fish, and the children were asked an appearance question: “When you look at this fish right now, does the fish look white or does it look blue?” (blue), and a reality question: “What colour is the fish really and truly, blue or white?” (white).

*Skewer in glass* (Gopnik & Astington, 1988). The children were shown a skewer and told, “See this stick. I’m going to place it in a glass of water”. The skewer was then placed in a glass of water touching the bottom at a 45-degree angle, making it appear bent. The children were then asked an appearance question: “When you look at this with your eyes right now, does this stick look bent and crooked or does it look straight?”
(bent), and a reality question: “What is this stick really and truly? Is it straight or is it bent and crooked?” (straight).

Crayon Box (Frye et al., 1995). The children were shown a crayon box. It was assumed that they would initially represent the contents incorrectly, by thinking there would be crayons inside. They were then shown that the box actually contained sticks and were asked, “What’s in the box?” The sticks were then put back in the box and the lid was closed. The children were asked an appearance question: “When you look at this box with your eyes right now, what does this box look like it has in it, sticks or crayons?” (crayons), and a reality question: “What’s really and truly in this box, sticks or crayons?” (sticks)

Smarties Box (Perner et al., 1987). Children were shown a smarties box, and it was assumed that they would initially represent the contents incorrectly. They were then shown the actual contents of the box and were asked to identify them (pencils). The box was then closed, and the children were asked a belief question: “If someone came into this room right now and had not seen what was inside the box, what would they think was in it, smarties or pencils?” (smarties), and a reality question “What is really and truly in the box, pencils or smarties?” (pencils).

Colour (Gopnik & Astington, 1988). The children were shown a white cat covered by green transparent plastic, so that the cat appeared green. It was assumed that the children would initially represent the cat as green. The plastic was removed, revealing the cat to be white. The cat was then returned to its original deceptive state, and the children were asked a belief question: “When you first saw the cat, all covered up like this, what colour did you think the cat really was, green or white?” (green), and a reality question: “What colour is the cat really and truly, green or white?” (white).

Chocolate hiding task (Wimmer & Perner, 1983). Children were told a story about a boy (Maxi) who went shopping with his mother and bought some chocolate.
When they arrived home Maxi put the chocolate into the blue cupboard (matchbox) and went outside to play. While Maxi was away his mother moved the chocolate from the blue cupboard to the green cupboard (matchbox). She then went to talk to the neighbour. Maxi came in and wanted some chocolate. He remembered where he had put the chocolate. Belief and reality questions were, respectively: “Where will Maxi look first for the chocolate, in the blue cupboard or in the green cupboard?” (blue), and “Where is the chocolate really and truly, in the blue cupboard or in the green cupboard?” (green).

Each appearance-reality and false-belief item was scored as correct if responses to the component belief/appearance and reality questions were both correct. This scoring reflects the complexity analysis, that is, children must be able to create an integrated representation to pass the ternary level items. Thus, the maximum score for the theory-of-mind items (appearance-reality and false-belief combined) was 6.

Results

Initial examination of the data indicated that the perfect performance of 5- and 6-year-old children on the binary-relational connections and transformation items produced zero variance for these cells. This resulted in non-normal distributions for this age group on these tasks, and non-equal error variance across all conditions. Consequently, the assumptions of ANOVA were not met, and therefore non-parametric techniques were used to evaluate predictions regarding differences in mean levels of performance between age groups (Kruskall-Wallis), and between tasks (Wilcoxon signed-rank test).

Complexity Effects

A Wilcoxon signed-rank test was conducted to compare performance on the connections/transformation items and the appearance-reality/false-belief items. Due to the different number of items in each level of complexity, scores for the two levels of
complexity were converted to percentages, and are presented in Figure 7.2. For the entire sample \( (N = 120) \), a significant difference was found between the children’s performance on the connections/transformation items and appearance-reality/false-belief items, \( Z = -7.07, p < .001 \). Accuracy was higher for the binary-relational connections/transformation items than for the ternary-relational appearance-reality/false-belief items. Separate analyses conducted for each of the four age groups revealed significant differences for the 3-year-olds, \( Z = -5.29, p < .001 \), and 4-year-olds, \( Z = -4.24, p < .001 \). There was no significant difference in performance of the two levels of complexity for 5-year-olds, \( Z = 1.73, p = .083 \), or 6-year-olds, \( Z = -1.00, p = .317 \).

![Figure 7.2](image)

*Figure 7.2.* Mean (standard error) percentage of items correct for each age-group on the connections/transformation items and appearance-reality/false-belief items \( (N = 120) \).

**Age Effects**

The connections/transformation items and the appearance-reality/false-belief items were analysed separately to investigate the effect of age on performance. The non-parametric Kruskall-Wallis test was used. A significant effect of age was observed on the connections/transformation items, \( \chi^2(3, N = 120) = 65.87, p < .001 \), with means ranks of 27.77 for 3-year-olds, 57.23 for 4-year-olds, and 78.5 for both 5- and 6-year-old children.
olds. A significant effect of age was also observed for the appearance-reality/false-belief items, $\chi^2(3, N = 120) = 83.82, p < .001$, with mean ranks of 22.35 for 3-year-olds, 45.38 for 4-year-olds, 81.12 for 5-year-olds and 93.15 for 6-year-olds. Performance on the theory-of-mind tasks at both levels of complexity increased with age.

**Performance in Relation to Chance**

Single sample $t$ tests were used to compare the means for 3- and 4-year-olds on the binary-relational and ternary-relational items, and the means for the 5- and 6-year-olds on the ternary-relational items, to their respective chance levels. The errorless performance of the 5- and 6-year-olds on the binary-relational items precluded $t$-tests for those means. Due to the joint scoring method used for the tasks, the chance level was 0.25. Thus chance corresponded to 2 out of 8 for the binary-relational items, and 1.5 out of 6 for the ternary-relational items. For the binary-relational versions, the mean scores of 3-, and 4-year-olds were both significantly above chance, smallest $t(29) = 22.65, p < .001$. For the ternary-relational version, the mean score of 3-year-olds was significantly below chance, $t(29) = -2.77, p = .010$. The mean scores of 4-, 5- and 6-year-olds were significantly higher than chance, smallest $t(29) = 3.81, p = .001$.

Performances of the individual children were evaluated against chance. Binomial tables indicated that scores of 5 or more out of 8 were significantly above chance on the binary-relational items, while scores of 4 or more out of 6 were significantly above chance on the ternary-relational items ($p < .05$). As can be seen in Table 7.1, the majority of all age groups performed significantly above chance level on the binary-relational items. In contrast, a minority of 3- and 4-year-olds and a majority of 5- and 6-year-olds performed significantly above chance on the ternary-relational items. Further, the pass-fail data indicated that the median age of attainment is three years (or earlier) for the binary-relational items, but not before five years for the ternary-relational items.
Table 7.1

*Frequency of Children within Each Age Group Performing above Chance Levels on the Binary-relational and Ternary-relational Theory-of-Mind items (N = 120)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Connections/Transformations</th>
<th>Theory-of-Mind</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>4-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>5-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>6-years</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 7.2 relates the pass-fail scores for the binary-relational connections and transformation items and the ternary-relational theory-of-mind items. Fifty-eight percent of children (70) performed consistently, failing or passing items at both levels of complexity. The remaining (50) children showed an inconsistent pattern. These children were significantly more likely to pass the connections and transformation items and fail the theory-of-mind items (100%) than to show the reverse pattern, McNemar test, $\hat{\chi}^2(N = 120, 1) = 48.02, p < .001$.

Table 7.2

*Crosstabulation of Children Passing and Failing the Binary- and Ternary-Relational Theory-of-Mind Items (N = 120)*

<table>
<thead>
<tr>
<th>Connections/Transformation</th>
<th>Theory-of-Mind</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
</tr>
<tr>
<td>Fail</td>
<td>2</td>
</tr>
<tr>
<td>Pass</td>
<td>50</td>
</tr>
</tbody>
</table>
Discussion

Findings of the current study are consistent with the relational complexity interpretation of children’s difficulty on theory-of-mind tasks. All age groups performed significantly above chance on the binary-relational connections and transformation items, with 96% of 3-year-olds passing. In contrast, only 5- and 6-year-old children were able to successfully complete the ternary-relational false-belief and appearance-reality items. Thus, whereas 3- and 4-year-olds are able to process the component binary relations involved in these tasks, they appear to be unable to integrate these binary relations into the ternary relation required for successful completion of the false-belief and appearance-reality items.

The findings of the current study are consistent with previous research. For example, Andrews et al. (2003, Exp. 3) found that the majority of 3-, 4-, and 5-year-old children performed significantly above chance on the binary-relational items. In contrast only 3.57% of 3-year-olds, 35.7% of 4-year-olds and 75% of 5-year-olds demonstrated theory-of-mind ability on the ternary-relational items.

The results can also be accounted for by Frye et al.’s (1995) complexity explanation. These authors suggested that advances in children’s performance on theory-of-mind tasks reflected improvement in their reasoning based on embedded rules. That is, to succeed on the more complex theory-of-mind items, children need to not only consider their own mental states and perspectives, but also to consider those of another. This requires the child to use embedded rules in order to switch between judgments based on these two conflicting perspectives, a skill that may not develop until 4- to 5-years of age (Frye et al., 1995).

The performance of children on the connections and transformation items is consistent with the results of Flavell et al. (1990), who reported that 75-80% of 3-year-old children passed the perspective-taking questions in their connections task. The
difficulty experienced by the 3-year-old children in the current sample on the ternary-relational theory-of-mind tasks is also consistent with prior research. For example, the meta-analysis conducted by Wellman et al. (2001) indicated that children did not perform above chance levels on false-belief tasks until four years of age or older. Similarly, Flavell et al. (1983) also found that 4-year-olds were much better at making the appearance-reality distinction than 3-year-olds, and that 5-year-olds performed almost errorlessly. Only 36% of 4-year-old children in the current sample performed above chance on the ternary-relational theory-of-mind items. This appears to be a transitional period with children’s understanding of appearance-reality and false-belief only beginning to develop. Thus, the current study confirmed the much replicated finding that theory-of-mind abilities emerge at around four to five years of age (Gopnik & Astington, 1988; Perner et al., 1987).

As described earlier, non-parametric statistics were utilized in the analysis of these tasks due to the non-equal error variance produced by the perfect performance of 5- and 6-year-olds on the connections and transformation items. These violations reflect data that are not inconsistent with the complexity theory. Differences in variance across age groups on the same task may reflect developmental changes. Periods of transition during which new understanding of concepts develops (e.g., 4-year-olds on theory-of-mind tasks) would be expected to produce greater variance than in younger age groups where all children fail, or older age groups where all children pass.

In summary, while all age groups performed significantly above chance on the connections and transformation items, only 5- and 6-year-old children were able to successfully complete the false-belief and appearance-reality items. The results indicate that the age-related improvements on theory-of-mind understanding are attributable to the complexity of the inferences required by the theory-of-mind tasks.
CHAPTER 8. COOL EXECUTIVE FUNCTION TASKS

Complexity analyses have been applied to a number of cool domains of executive function. The current research employed three measures that have previously been demonstrated to involve complexity in successful performance: (1) transitivity (e.g., Andrews & Halford, 2002); (2) class inclusion (e.g., Andrews & Halford, 2002); and (3) the DCCS (e.g., Zelazo & Frye, 1998).

Transitivity

An important characteristic of human behaviour is the ability to manipulate knowledge in order to deduce novel facts (Acuna, Eliassen, Donoghue & Sanes, 2002). One form of deductive reasoning is transitive inference, which is an important development milestone that Piaget attributed to the concrete operational stage of development that children progressed to at around seven to eight years of age.

Transitive inference involves ordinal relations (Halford, 1982). For example, given the premises, Tom is happier than Bob, Bob is happier than Mike, a child who understood transitivity could supply the appropriate answer (“Tom”) if asked “Who is happier, Tom or Mike?” Transitive reasoning is demonstrated when an inference $A R C$ is deduced from premises $A R B$ and $B R C$, where $R$ is a transitive relation, and $A$, $B$, and $C$ are the elements related. Determining the relation between $A$ and $C$ requires that premises $A R B$ and $B R C$ be integrated to form an ordered set of three elements, $A R b R C$. Premise integration relates three elements; therefore it is a ternary relation.

Andrews and Halford (2002) tested children’s ability to make transitive inference about the spatial position of coloured squares in a tower (based on a procedure developed by Pears & Bryant, 1990). Binary-relational items were also included. The study examined the performance of 3- to 8-year-old children on these binary-and ternary-relational transitivity items, as well as several other task. The results indicated a strong effect of relational complexity on the transitivity items, with the younger children...
mastering the binary-relational items, but experiencing difficulty on ternary-relational items. Significant correspondence was also found between transitivity and the other tasks presented at the same levels of complexity, including class inclusion, hierarchical classification, cardinality, comprehension of relative clause sentences, and hypothesis testing.

Transitivity is considered to be a cool executive function as it is relatively abstract and decontextualised (Andrews et al., 2003). Additionally, the DL-PFC has been implicated in the performance of reasoning tasks. For example, Kroger et al. (2000) recently employed an fMRI study to identify brain regions that responded selectively to processing high levels of relational complexity. The authors found significant activation of the DL-PFC during the performance of nonverbal reasoning problems, with the highest levels of complexity selectively activating left anterior regions. Acuna et al. (2002) utilized fMRI when participants performed transitive inference on an ordered list of 11 items. Significant activations were found in the dorsolateral portions of the superior, middle and inferior frontal gyri during the performance of transitive inference. Waltz et al. (1999) also examined the performance of a premise integration task, similar to transitivity, in patients with damage to the frontal cortex, including severe damage to the DL-PFC. The results indicated a severe deficit in the performance of the premise integration task by frontal patients, but normal levels of performance by patients with anterior temporal lobe damage. In contrast, the prefrontal patients performed more accurately than the temporal patients on tests of both episodic memory and semantic knowledge. The authors concluded that the neural system responsible for integrating multiple relations is sensitive to prefrontal damage, particularly the DL-PFC.

The current study administered binary- and ternary-relational transitivity items based on the design developed by Andrews and Halford (2002). Consistent with the
relational complexity analyses, it was anticipated that the majority of children in each age group would pass the binary-relational transitivity items. Only 5- and 6-year-olds were anticipated to pass the ternary-relational items.

Method

Materials. The same materials were used for ternary and binary items. Each premise display consisted of four pairs of coloured squares in which one colour is higher (nearer to the top of the page) than another (see Figure 8.1). The pairs together defined a unique vertical ordering of five coloured squares in a tower. For the example shown, the correct top-down order is yellow, green, red, blue, purple. More generally, \( A > B > C > D > E \), where \( A \) is top position and \( E \) is bottom. A different assignment of colours to ordinal positions was used on each trial, and the left-right order of the premises varied across displays. Additional coloured squares identical to those used in the premise displays were attached to 3-cm squares of cardboard. Children used these squares to construct their towers.

Procedure. Children were invited to play a game in which they would build towers using coloured squares. The importance of placing squares in the correct order was emphasised. Two practice trials were administered to familiarize children with using the premise information (clues) to determine the vertical spatial order of the coloured squares. Care was taken to ensure that they realised that the relation ‘higher than’ referred to adjacent squares (e.g., positions \( B \) and \( C \)) as well as to nonadjacent squares (e.g., positions \( B \) and \( D \)), and that a square could be inserted above, between, or below squares that were already in place. The experimenter reminded children to use the premise information to determine the correct order and also indicated that the children might need to look at more than one clue to determine the correct order.
In the binary-relational items, children were required to construct two towers, each containing five squares, beginning with an internal pair, either BC or CD. The experimenter gave the child two squares, which the child then attempted to place in the correct order. In Figure 8.1, BC is represented by green-red. To place green and red
correctly, children needed to use the premise pair, green above red. This is equivalent to a binary relation. Adding each subsequent square required consideration of a single premise. To place blue, children needed to refer to the premise, red above blue, then place blue below red, yielding the order green, red, blue. This was a concatenation procedure, which entailed processing one binary relation at a time. One point was awarded for each correctly ordered initial pair and subsequent square, yielding a maximum score of 8 (for 2 towers).

In the ternary-relational items, two squares corresponding to positions B and D were handed to the child, who then attempted to place the squares in the correct order (i.e., B above D). In Figure 8.1, B and D correspond to green and blue. Two premises – green above red and red above blue – must be integrated to form the ordered set, green above red above blue, from which green above blue can be deduced. As a check on guessing, red was placed after green and blue. If the child integrates BC and CD to conclude B above D, the correct position of C between B and D should be apparent. Eight items were administered. Credit was given for responses where B, D and C were placed correctly, so that participants receive 1 point for each correct ternary relation, yielding a maximum score of 8.

Results and Discussion

The number of correct responses on the transitivity task was analysed with a 4 age: 3, 4, 5, 6 years) × 2 (gender: male, female) × 2 (complexity: binary, ternary) ANOVA. The analysis indicated significant main effects of complexity, $F(1, 112) = 1421.81, p < .001, \eta^2 = .927$, and age, $F(3, 112) = 123.05, p < .001, \eta^2 = .767$. The main effects were moderated by a significant Complexity × Age interaction, $F(3, 112) = 60.15, p < .001, \eta^2 = .617$. Means and standard errors for the interaction are presented in Figure 8.2. Analysis of simple effects indicated that performance increased significantly across the four age groups on both the binary-relational, $F(3, 119) = 27.48, p < .001, \eta^2$
Development of Executive Functions

=.416, and ternary-relational items, \( F(3, 119) = 114.26, p < .001, \eta^2 = .747 \). The magnitude of this effect, however, was much stronger on the ternary-relational items than the binary-relational items. For the binary-relational items, a Scheffe test revealed significant differences between means for 3-year-olds and the 4-, 5-, and 6-year-olds only (all \( p \)'s < .001). For the ternary-relational items, a Scheffe test revealed significant differences between means for 3-year-olds and the 5- and 6-year-olds, 4-year-olds and 5- and 6-year-olds, and 5-year-olds and 6-year-olds (all \( p \)'s < .001). No main effect or interactions involving gender were recorded.

![Graph showing mean performance by age group](image)

**Figure 8.2.** Mean performance of the four age groups for the binary- and ternary-relational items of transitivity (\( N = 120 \)).

Single sample \( t \)-tests were used to compare the means for each age group to chance levels for the total binary-relational and ternary-relational scores. The chance level was 4 out of 8 for the binary-relational items, and 1.34 out of 8 for the ternary-relational items. For the binary-relational items, the mean scores of 3-, 4-, 5-, and 6-year-olds were all significantly above chance, smallest \( t(29) = 11.31, p < .001 \). For the ternary-relational items, the mean scores were significantly below chance levels for 3-
year-olds, \( t(29) = -11.57, p < .001 \) and 4-year-olds, \( t(29) = -13.94, p < .001 \). The mean scores were significantly above chance for 5-year-olds, \( t(29) = 6.52, p < .001 \), and 6-year-olds, \( t(29) = 13.30, p < .001 \).

Performances of the individual children were evaluated against chance. According to the binomial table, scores of 7 or more out of 8 were significantly above chance for the binary-relational items, while scores of 4 or more out of 8 were significantly above chance on the ternary-relational items. \( p < .05 \). As can be seen in Table 8.1, 53\% of 3-year-olds performed significantly above chance on the binary-relational items. The majority of 4-, 5- and 6-year-olds performed above chance levels on the binary-relational items. A minority of 3- and 4-year-olds, 50 per cent of 5-year-olds and a majority of 6-year-olds performed significantly above chance on the ternary-relational items. The pass-fail data indicated that the median age of attainment was five years for the ternary-relational items of the transitivity task, and three years for the binary-relational items.

Table 8.1

*Frequency of Children Within Each Age-Group Performing above Chance Levels on the Binary- and Ternary-Relational Items of Transitivity (N = 120)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Items</th>
<th>Ternary-Relational Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>4-years</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>5-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>6-years</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 8.2 relates the pass-fail scores for the binary- and ternary-relational items. Forty-eight percent of children (58) performed consistently, failing or passing items at both levels of complexity. The remaining (62) children showed an inconsistent pattern, passing the binary-relational-items and failing the ternary-relational items, McNemar test, $\chi^2(1, N = 120) = 60.02$, $p < .001$.

Table 8.2

*Crosstabulation of Children Passing and Failing the Binary- and Ternary-Relational Items of Transitivity (N = 120)*

<table>
<thead>
<tr>
<th></th>
<th>Ternary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>Binary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fail</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Pass</td>
<td>62</td>
<td>42</td>
</tr>
</tbody>
</table>

The transitivity results are consistent with the relational complexity approach. The majority of children in all age groups succeeded on the binary-relational transitivity items. In contrast, only 5- and 6-year-old children succeeded on the ternary-relational items. The findings of the current study are consistent with previous complexity research. For example, Andrews and Halford (2002) also found that the ternary-relational items of transitivity were not passed until at least five years of age.
Development of Executive Functions

Class Inclusion

An additional core cognitive process that develops during the preschool years is the ability to reason about hierarchically related classes. One measure utilized by Piaget to assess this ability is the class inclusion paradigm. This task utilizes items in which pictures of two different subclasses (e.g., 18 apples and 2 pears) belonging to the same superordinate class (e.g., fruit) are presented. Children were asked a class inclusion question in which the larger subclass was compared to its superordinate class. Piaget found that children between three and seven years usually used subclass-subclass reasoning in attempting to solve the question. For example, in the class inclusion question “Are there more apples or more fruit?”, young children treat “fruit” as “pears” and incorrectly answered “more apples”. In contrast, older children usually solved the question using the appropriate class-inclusion reasoning.

According to Piaget, success on class inclusion tasks depends on the child’s ability to perform two operations, one involving the addition of classes (e.g., apples + pears = fruit) and the other involving the subtraction of classes (e.g., apples = fruit – pears). Class inclusion involved conserving the whole while maintaining the identity of the parts, that is children must be able to conserve the superordinate class (fruit) while maintaining each of the subclasses (apples, pears; Winer, 1980).

According to Halford (1982), class inclusion tasks entail presenting a superordinate set with two nonempty subsets. This is a ternary relation among three classes: inclusion (fruit, apples, nonapple-fruit). There are also three binary relations: inclusion (fruit, apples), inclusion (fruit, nonapple fruit) and complement (apples, nonapple-fruit). Notice that no single binary relation is sufficient for understanding inclusion. For example, it is not enough to know that apples are related to fruit because, unless the relation between fruit and nonapple fruit is also considered, it is not possible to distinguish inclusion from complete overlap (Johnson, Scott & Mervis, 1997). Unless
it is recognized that nonapple fruit is also fruit, it is not possible to decide whether fruit and apples are coextensive. Therefore, the inclusion hierarchy cannot be decomposed into a set of binary relations without losing the essence of the concept (Andrews & Halford, 2002). Binary level understanding is assessed through comparison of the two sub-sets, for example, by asking, “Are there more apples or more pears?”

Consistent with the relational complexity interpretation, children as young as five years of age have been found to succeed on the class inclusion task. The aforementioned study by Andrews and Halford (2002) on 3- to 8-year-old children also included a class inclusion task. A strong effect of relational complexity was seen, with the younger children mastering the binary-relational class inclusion items, but experiencing difficulty on ternary-relational items. It was not until approximately five years of age that children were able to succeed on the ternary-relational items of the class inclusion task.

The class inclusion task can be seen to parallel other measures of cool executive functioning. For example, it requires the consideration of category membership in a similar manner to the way cards are sorted in the WCST. It is also a relatively abstract and decontextualised measure, that involves no motivation or reward based contingencies. For these reasons, the class inclusion task was included as a measure of cool executive function in the current study. The current study administered binary- and ternary-relational class inclusion items. Consistent with the relational complexity analyses, it was anticipated that the majority of children in each age group would be able to succeed on the binary-relational transitivity items. Only 5- and 6-year-olds were anticipated to pass the ternary-relational items.
Method

Materials. Six stimulus sets were used which consisted of geometric shapes cut from coloured paper and attached to A4-sized pieces of paper. A typical display contained three yellow squares and two blue squares. The composition of the displays, which were identical to Hodkin’s (1987), is shown in Table 8.3. The shapes in each set were arranged in a single row across the card (landscape orientation).

Table 8.3

Composition of Displays used in Class Inclusion Task

<table>
<thead>
<tr>
<th>Set Number</th>
<th>Superordinate</th>
<th>Major Subclass</th>
<th>Minor Subclass</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Circles</td>
<td>4 green circles</td>
<td>3 yellow circles</td>
</tr>
<tr>
<td>2</td>
<td>Triangle</td>
<td>3 blue triangles</td>
<td>2 red triangles</td>
</tr>
<tr>
<td>3</td>
<td>Squares</td>
<td>3 yellow squares</td>
<td>2 blue squares</td>
</tr>
<tr>
<td>4</td>
<td>Green</td>
<td>3 green squares</td>
<td>1 green circle</td>
</tr>
<tr>
<td>5</td>
<td>Red</td>
<td>4 red triangles</td>
<td>2 red squares</td>
</tr>
<tr>
<td>6</td>
<td>Blue</td>
<td>5 red circles</td>
<td>2 blue triangles</td>
</tr>
</tbody>
</table>

Procedure. The children were introduced to the task with the sentence, “Now we are going to talk about some colours and shapes”. The procedure will be demonstrated using set 3 in Table 8.3. The experimenter identified the classes by saying, “These are the blue things, these are the yellow things, and these are the squares”, touching the appropriate elements of the display slowly while making the statements. Children were then asked three questions per display: (1) Subclass comparison: “Are there more yellow things or more blue things?” (2) Superordinate class-major subclass comparison:
“Are there more squares or more yellow things?” and (3) Superordinate class-minor subclass comparison: “Are there more squares or more blue things?” An analogous procedure was followed for each of the five displays. Responses were recorded by the experimenter on a response sheet (see Appendix G).

Two orders of presentation were used within each age group. Binary relation scores (out of 6) were based on correct responses to sub-class comparison questions. Ternary relation scores were derived from the superordinate-major subclass, and superordinate class-minor subclass comparisons. Guessing was estimated from errors on the superordinate class-minor subclass questions based on Hodkin’s (1987) rationale and procedure. Thus, the ternary relation score (maximum of 6) reflected correct responses to the superordinate class-major sub-class question minus errors on the superordinate class-minor subclass questions (minimum = 0). This yielded a maximum score of 6.

**Results and Discussion**

The number of correct items on the class inclusion task was analysed using a 4 age: 3, 4, 5, 6 years) × 2 (gender: male, female) × 2 (complexity: binary, ternary) ANOVA. The analysis indicated significant main effects of complexity, $F(1, 112) = 452.43, p < .001, \eta^2 = .802$, and age, $F(3, 112) = 49.10, p < .001, \eta^2 = .674$. The main effects were moderated by a significant Complexity × Age interaction, $F(3, 112) = 77.17, p < .001, \eta^2 = .568$. Means and standard errors for the interaction are presented in Figure 8.3. The negative result that can be seen for 3-year-olds can be attributable to the scoring procedure described above that involved adjustments for guessing. Analysis of simple effects indicated that the performance of children increased significantly across the four age groups on both the binary-relational, $F(3, 119) = 22.93, p < .001, \eta^2 = .372$, and ternary-relational items, $F(3, 119) = 70.83, p < .001, \eta^2 = .647$. The magnitude of this effect however, was stronger on the ternary-relational items than the binary-
relational items. For the binary-relational items, a Scheffe test revealed significant differences between means for 3-year-olds and the 4-, 5-, and 6-year-olds only (all \( p \)’s < .001). For the ternary-relational items, a Scheffe test revealed significant differences between means for 3-year-olds and the 5- and 6-year-olds, and 4-year-olds and 5- and 6-year-olds (all \( p \)’s < .001). No main effect or interactions involving gender were recorded.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure83.png}
\caption{Mean performance of the four age groups across the binary- and ternary-relational items of the class inclusion task (\( N = 120 \)).}
\end{figure}

Single sample \( t \)-tests were used to compare the means for each age group to their respective chance levels for the binary-relational and ternary-relational scores. The chance level was 3 out of 6 for the binary-relational items, and 1.5 out of 6 for the ternary-relational items. For the binary-relational items, the mean scores of 3-, 4-, 5-, and 6-year-olds were all significantly above chance, smallest \( t(29) = 9.00, p < .001 \). For the ternary-relational items, the mean scores were significantly below chance levels for 3-year-olds, \( t(29) = -5.41, p < .001 \) and 4-year-olds, \( t(29) = -4.95, p < .001 \). The mean
scores were significantly above chance for 5-year-olds, $t(29) = 8.76, p < .001$, and 6-year-olds, $t(29) = 12.30, p < .001$.

Performances of the individual children were evaluated against chance. According to the binomial table, scores of 6 out of 6 on the binary-relational items and 4 or more out of 6 on the ternary-relational items are significantly above chance ($p < .05$). As can be seen in Table 8.4, a minority of 3-year-olds, and a majority of 4-, 5- and 6-year-old children performed above chance levels on the binary-relational items. When a less stringent criterion (5/6) is used on the binary-relational items, a majority of children in all age groups performed above chance. A minority of 3- and 4-year-olds and a majority of 5- and 6-year-olds performed significantly above chance on the ternary-relational items. The pass-fail data indicated that the median age of attainment was four years for the binary-relational items of the class inclusion task, and not until five years for the ternary-relational items.

Table 8.4

*Frequency of Children Within Each Age-Group Performing above Chance Levels on the Binary- and Ternary-Relational Items of Class Inclusion (N =120)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Items</th>
<th>Ternary-Relational Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>4-years</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td>5-years</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>6-years</td>
<td>1</td>
<td>29</td>
</tr>
</tbody>
</table>
Table 8.5 relates the pass-fail scores for the binary- and ternary-relational items of the class inclusion task. Seventy percent of children (84) performed consistently. The remaining children (36) showed an inconsistent pattern. These children were significantly more likely to pass the binary-relational items and fail the ternary-relational items of class inclusion (97.22%) than to show the reverse pattern (2.77%), McNemar test, $\chi^2(N = 120, 1) = 30.25, p < .001$.

Table 8.5

*Crosstabulation of Children Passing and Failing the Binary-Relational and Ternary-Relational Items of Class Inclusion (N = 120).*

<table>
<thead>
<tr>
<th></th>
<th>Ternary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>Fail</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>Pass</td>
<td>35</td>
<td>52</td>
</tr>
</tbody>
</table>

The results of class inclusion obtained in the current study are consistent with the relational complexity approach. The majority of children in all age groups succeeded on the binary-relational class inclusion items. In contrast, only 5- and 6-year-old children succeeded on the ternary-relational items. These findings are also consistent with previous complexity research, which found that the ternary-relational items of class inclusion were not passed until at least five years of age (Andrews & Halford, 2002).
Dimensions Change Card Sorting Task

Cool executive functions can be further examined using a representational rule use task. As mentioned previously, the DCCS involves children being shown two target cards (e.g., a red circle and a green triangle) that remain visible throughout the task, and children are asked to sort test cards (e.g., green circles and red triangles) that match one of the target cards on one dimension (e.g., shape) and match the other target card on the other dimension (e.g., colour). Children are first told explicitly to sort test cards by one dimension for a number of trials and then they are asked to switch and sort cards by the other dimension.

According to the Cognitive Complexity and Control theory (Zelazo & Frye, 1998) the DCCS involves a rule for assigning shapes, a rule for assigning colours, and a higher-order rule that switches between shape and colour settings. Simpler tasks entail only the shape or colour rules, so complexity is defined by the number of levels of the hierarchical rules. Using the DCCS, Frye et al (1995) found that regardless of which sorting dimension was presented first, 3-year-olds typically perseverated by continuing to sort cards by the first dimension only. In comparison, 4- and 5-year old children performed well. They successfully switched between sorting dimensions.

Zelazo et al. (2003) conducted a series of nine experiments that utilized standard and modified versions of the DCCS. The authors argued against several proposed explanations of young children’s difficulty on the task, such as working memory limitations, representational redescription, and inhibition of particular stimulus configurations or the preswitch dimensions. The findings indicated that errors occur because the sorting rules are activated in the preswitch phase and this activation persists during the postswitch phase. Some evidence was also found for the negative priming account, which attributes children’s difficulty in the postswitch phase to their failure to disinhibit the previously irrelevant rule. However, negative priming occurs only when
children must actively select one pair of rules against a competing alternative, that is, when there is a conflicting mismatch between target and test cards during the postswitch phase. Thus, the authors concluded that a key determinant of children’s difficulty across all studies was the ability to shift between conflicting rule-pairs.

According to relational complexity theory, it is the inability to decompose the DCCS into simpler subtasks that is the cause of the cognitive complexity. The DCCS task cannot be decomposed into two subtasks that are performed independently, because the conflicting dimension is always present. That is, when more than one game is being played, the setting condition has to be processed along with the colour or shape attributes. For example, children must remember that they are playing the colour game in order to determine that red triangle is sorted with the red circle, and not the green triangle (see Figure 8.4, standard task). Thus, the DCCS in the original format involves the relation between three variables: the setting condition (shape, colour), sorting card attribute, and the sorting category. The task is therefore ternary-relational.

The binary-relational task, in contrast, involves items that are decomposable into two subtasks in which the irrelevant dimension is held constant. For example, as can be seen in Figure 8.4 (decomposable stimuli), when sorting by the shape dimension in the binary-relational items, both shapes (triangle, circle) are presented in the same colour (red). When sorting by the colour rule, the two colours (green, red) are presented on the same shape (circle). These binary-relational items remove the conflict between shape and colour dimensions that occurs within each subtask. The setting condition can therefore be processed in advance, and each subtask requires sorting by only one dimension while the other is held constant. The binary-relational items have a similar hierarchical structure to the ternary-relational items because both shape and colour games are relevant as the items require switches between the games. The materials and procedures are also similar across the two levels of complexity.
The DCCS is considered a cool executive function task, as it is a refined version of the WCST in which children are explicitly told the sorting criterion. As outlined in Chapter 1, the DL-PFC has been found to be important for set-shifting and the ability to alternate between various categories, thereby suggesting a role of the DL-PFC in the performance of the DCCS also.

The current study administered both binary- and ternary-relational items of the DCCS. Consistent with the relational complexity analyses, it was anticipated that the majority of children in each age group would be able to succeed on the binary-relational DCCS items. Only 5- and 6-year-olds were anticipated to pass the standard ternary-relational items.

Figure 8.4. Stimuli for the Dimensional Change Card Sort task.
Method

Materials. Eight templates similar to those used by Frye et al., (1995) were constructed, each from an A4-sized piece of cardboard. The templates were created by dividing each piece of cardboard into half by either a vertical (templates 1a, 1b, 3a, 3b) or a horizontal line (templates 2a, 2b, 4a, 4b). Standards were attached to the tops of the columns or at the left of the rows. Templates 1a and 1b were identical except that the positions of the standards were reversed. Similarly, templates 2a, 3a, and 4a were identical to templates 2b, 3b, and 4b, respectively, except for the left-right or top-bottom positions of the standards. Standards were line drawings of the shapes (triangles, circles, flowers, boats) photocopied onto paper of the appropriate colour (red, green, blue, yellow), see Figure 8.5 for an example.

![Example template and objects sorted in DCCS task](image)

Figure 8.5. Example template and objects sorted in DCCS task

Procedure. As described by Andrews et al. (2003), the DCCS procedure was modeled on Frye et al (1995, Experiment 3), and will be explained with reference Template 1a which is shown in Figure 8.5. The template was placed in front of the child, and they were told “This is the colour (shape) game”. The experimenter placed
the red triangle explaining why it belongs in the column below the red circle, “See, I have a red one. It goes here because this one is red too”. The red triangle was then removed. Placement of the green circle in the column below the green triangle was then demonstrated, “See, I have a green one. It goes here because this one is green too”. Template 1b was presented and the child was asked to sort the same two objects (red triangle, green circle), “Here is a red one, put it in the space where it belongs”. The first object (red triangle) was removed before the second object (green circle) was sorted. Feedback was provided for sortings based on the first template presented: “Yes, that’s right” or “No, that’s not right”, followed by “This one goes here because it is green (red) like this one,” for sortings based on colour or, “This one goes here because it is circle (triangle) like this one,” for sortings based on shape.

Each child was administered four ternary-relational colour trials and four ternary relational shape trials. Two presentation orders were utilized, both of which involved switching between the two sorting dimensions. For example, the first presentation order followed the following pattern: colour, colour, shape, shape, shape, shape, colour, colour. The second presentation order followed the same pattern, however, it began with the shape trials (see Appendix H for information regarding the templates and objects sorted in the ternary-relational items). One mark was awarded for each object correctly placed, yielding a maximum score of 8. Response sheets are presented in Appendix I.

Binary level items were presented using similar task procedures. The sorting objects differed, however, as the irrelevant dimension was held constant. For example, when sorting by the shape dimension, both shapes were the same colour, or when sorting by the colour rule, the colours were presented on the same shape. Eight trials were presented (four colour trials, four shape trials, in a counterbalanced order), following the same pattern as the ternary relational items (see Appendix H). One mark awarded for each correctly placed item (maximum = 8).
Results and Discussion

The number of items correct on the DCCS was examined using a $4 \times 2 \times 2$ (age: 3, 4, 5, 6 years) ANOVA. The analysis indicated significant main effects of complexity, $F(1, 112) = 156.65, p < .001, \eta^2 = .588$, and age, $F(3, 112) = 49.35, p < .001, \eta^2 = .529$. The main effects were moderated by a significant Complexity $\times$ Age interaction, $F(3, 112) = 41.95, p < .001, \eta^2 = .569$. Means and standard errors for the interaction are presented in Figure 8.6.

![Figure 8.6](image)

*Figure 8.6. Mean performance of the four age groups for the binary- and ternary-relational items of the DCCS (N = 120).*

Analysis of simple effects indicated that performance increased significantly across the four age groups on both the binary-relational, $F (3, 119) = 8.45, p < .001, \eta^2 = .179$, and ternary-relational items, $F (3, 119) = 51.71, p < .001, \eta^2 = .572$. The magnitude of this effect, however, was much stronger on the ternary-relational items than the binary-relational items. For the binary-relational items, a Scheffe test revealed significant differences between means for 3-year-olds and the 4-year-olds ($p = .017$), and 3-year-olds and 5- and 6-year-olds only ($p$'s = .001). For the ternary-relational items, a Scheffe test revealed significant differences between means for 3-year-olds and
the 5- and 6-year-olds, and 4-year-olds and 5- and 6-year-olds (all $p$’s < .001). No main effect or interactions involving gender were recorded.

Single sample $t$-tests were used to compare the means for each age group to their respective chance levels for the total binary-relational and ternary-relational scores. The chance level was 4 out of 8 for both levels of complexity. The mean scores of 3-, 4-, 5-, and 6-year-olds were all significantly above chance on the binary-relational items, smallest $t(29) = 34.37, p < .001$ and ternary-relational items, smallest $t(29) = 6.01, p < .001$.

Performances of the individual children were evaluated against chance. The probability of a correct response on each item was 0.50, thus scores of 7 or more out of 8 are significantly above chance according to the binomial table ($p < .05$). As can be seen in Table 8.6, the majority of all age groups performed at above chance levels on the binary-relational items. A minority of 3- and 4-year-olds, and a majority of 5- and 6-year-olds performed significantly above chance on the ternary-relational items. The pass-fail data indicated that the median age of attainment was five years for the ternary-relational items of the DCCS task, and three years (or earlier) for the binary-relational items.

Table 8.6

*Frequency of Children in Each Age-Group Performing above Chance Levels on the Binary- and Ternary-Relational Items of the DCCS ($N = 120$)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Binary-Relational Items</th>
<th>Ternary Relational Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3-years</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>4-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>5-years</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>6-years</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 8.7 relates the pass-fail scores for the binary- and ternary-relational items of the DCCS task. Fifty-five percent of children (66) performed consistently, failing or passing items at both complexity levels. The remaining children (54) showed an inconsistent pattern. These children were significantly more likely to pass the binary-relational items and fail the ternary-relational items (100%) than to show the reverse pattern, McNemar test, $\chi^2(1, N = 120) = 45.02, p < .001$.

Table 8.7

*Crosstabulation of Children Passing and Failing the Binary- and Ternary-Relational Items of the DCCS (N = 120)*

<table>
<thead>
<tr>
<th></th>
<th>Ternary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail</td>
</tr>
<tr>
<td>Binary</td>
<td></td>
</tr>
<tr>
<td>Fail</td>
<td>1</td>
</tr>
<tr>
<td>Pass</td>
<td>54</td>
</tr>
</tbody>
</table>

The above results for the DCCS are consistent with relational complexity theory. The majority of all age groups succeeded on the binary-relational findings. In contrast, the majority of only the 5- and 6-year-olds succeeded on the ternary-relational items. Frye et al. (1995) also found that 3-year-olds were unable to switch to a new sorting criterion on the DCCS. Only, 4- and 5-year-old successfully switched between sorting dimensions. Frye et al. (1995) interpreted their results in terms of the CCC theory, suggesting that it was not until four to five years of age that children can typically represent a higher order rule that allows them to select between two incompatible pairs of rules. According to relational complexity theory, what 3- and 4-year-olds cannot do is process the integrated rule. However, they can process the two tasks in succession.
General Discussion

The aforementioned results of transitivity, class inclusion and the DCCS are consistent with the much replicated finding of significant development of cool executive functions during the preschool years. As described in Chapter 2, many cool executive functions emerge during the period of three to six years of age. However, the relation between the development of these cool executive functions associated with the DL-PFC and hot executive functions associated with the OFC remains unclear. The relation between executive abilities associated with these two domains will now be examined further in Chapter 9.
CHAPTER 9. CROSS TASK COMPARISONS

This chapter will compare the performance of children across the two executive function domains to determine whether functions associated with the OFC emerge earlier than those associated with the DL-PFC. This was examined in several ways. Firstly, the results from the analyses reported in the previous chapters were brought together to allow for a comparison of complexity and age effects across the tasks. This was done for analyses based on both the mean levels of performance and the pass-fail scores. Secondly, correlational analyses were conducted to determine the within and across domain associations. Finally, data reduction techniques were applied. Factor scores for hot and cool executive function tasks were computed and the effects of age and domain were examined.

**Analyses based on Mean Performance**

The preceding chapters (4 to 8) described the results for each task at the two levels of complexity. A general finding was that binary-relational items were easier than the ternary-relational items for each task. The mean scores for binary-relational items on all tasks were significantly above chance for even the 3-year-old participants. In contrast, the mean scores for the ternary-relational items were typically above chance for the 5- and 6-year-olds only. However, variability was evident across the different tasks, particularly in the performance of 4- and 5-year-olds. This was examined by comparing the patterns of age-related improvement revealed by the Scheffe tests reported in previous chapters. The significance of the improvements between the mean performance of 3- and 4-year-olds, 4- and 5-year-olds, and 5- and 6-year-olds on the ternary-relational items are presented in Table 9.1. This table reveals significant improvement between the means of 3-year-olds and 4-year-olds on all four of the hot measures (gambling task, future-oriented, conditional discrimination and theory of mind). In contrast, there was no significant improvement between the means of 3- and
4-year-olds on the three cool tasks (transitivity, class inclusion, DCCS). Significant improvements between 4-year-olds and 5-year-olds were found for all seven executive tasks. This pattern suggests that hot executive functions begin to emerge slightly earlier in development than cool executive functions.

Table 9.1

*Patterns of Age-Related Changed revealed by Scheffe tests on the Ternary-Relational Items (N = 120).*

<table>
<thead>
<tr>
<th>Task</th>
<th>Significance Levels Between Means of Age Groups:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3- and 4-years</td>
</tr>
<tr>
<td>Gambling</td>
<td>( p &lt; .001 )</td>
</tr>
<tr>
<td>Future-oriented</td>
<td>( p = .001 )</td>
</tr>
<tr>
<td>Conditional Discrimination</td>
<td>( p = .005 )</td>
</tr>
<tr>
<td>Theory-of-mind</td>
<td>( p &lt; .001 )</td>
</tr>
<tr>
<td>Transitivity</td>
<td>( ns )</td>
</tr>
<tr>
<td>Class inclusion</td>
<td>( ns )</td>
</tr>
<tr>
<td>DCCS</td>
<td>( ns )</td>
</tr>
</tbody>
</table>

*Analyses based on Percentages of Children Passing the Tasks*

The variability of performance across the hot and cool ternary-relational items was further examined by investigating the number of children who passed these items. Table 9.2 shows the percentage of children in each group whose performance on the tasks indicated they were able to process ternary relations. Data for the gambling,
future-oriented and conditional discrimination tasks were combined across the pilot and main studies.

Table 9.2

*Percentage of Children in Each Age Group Passing Ternary-Relational Items*

<table>
<thead>
<tr>
<th>Task</th>
<th>Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Gambling task</td>
<td>14.81</td>
</tr>
<tr>
<td>Future-oriented</td>
<td>0.00</td>
</tr>
<tr>
<td>Conditional Discrimination</td>
<td>7.41</td>
</tr>
<tr>
<td>Theory-of-mind</td>
<td>3.33</td>
</tr>
<tr>
<td>Transitivity</td>
<td>0.00</td>
</tr>
<tr>
<td>Class inclusion</td>
<td>3.33</td>
</tr>
<tr>
<td>DCCS</td>
<td>10.00</td>
</tr>
</tbody>
</table>

Note: Gambling, Future-oriented decision making and Conditional Discrimination samples include pilot data thus $n = 54$; For the remaining tasks $n = 30$.

Examination of Table 9.2 indicates a steady increase in the percentage of children passing the ternary-relational items from three to four to five to six years of age. For all seven tasks the median age of attainment was five years. There was little variability in the performance of the 3-year-old children across the tasks, with the majority of children (85-100%) failing these items. There was also little variability in the performance of 6-year-old children across the seven tasks, with the majority of these children (77-100%) passing the ternary-relational items. However, significant variation
was evident in the performance of 4-year-olds and 5-year-olds. Firstly, three tasks from
the hot domain (gambling, conditional discrimination, theory-of-mind) showed a similar
pattern of emergence. Between 37 and 39% of 4-year-olds and at least 90% of 5-year-
olds passed these hot tasks. The three cool measures (transitivity, class inclusion,
DCCS) also showed a similar pattern to each other with 0 to 17% of 4-year-olds and 50
to 84% of 5-year-olds passing these cool ternary-relational items. Thus, consistent with
the Scheffe tests (Table 9.1) success on the ternary-relational items of three of the hot
executive function tasks appeared to emerge earlier than success on the ternary-
relational items of the cool executive function tasks.

The pass-fail performance of 4- and 5-year-olds on the future-oriented decision-
making task revealed a divergent pattern relative to the other hot measures. No 3-year-
olds and only 13% of 4-year-olds passed this task. The observed pattern is more
consistent with the cool tasks. This is in contrast to the pattern revealed by the age-
related improvements (Scheffe’s tests) where the future-oriented decision-making task
demonstrated the same pattern as the other hot executive function tasks. Thus, an
inconsistent pattern was revealed regarding this measure. As mentioned earlier,
however, the involvement of episodic representation in the performance of future-
oriented decision-making might cast doubt over its ‘hotness’ and the associated
involvement of the OFC.

The previous two analyses were repeated using only those children who had
passed the binary-relational items for each task. These analyses revealed no significant
change in terms of the patterns that emerged, therefore the data of all children was
retained.
Intercorrelations Among Tasks

Correlational analyses were then conducted to examine the strength of associations of tasks within and across domains. Based on the hot-cool distinction, significant correlations were expected between tasks from the same domain. This would be consistent with previous research that has found correlations between hot tasks (e.g., Moore et al., 1998) and correlations between cool tasks (e.g., Andrews & Halford, 2002). Correlations were also expected across the two (hot and cool) domains, as Andrews et al. (2003) found associations between a hot task (theory-of-mind) and several cool tasks.

Pilot Study

For each hot task included in the pilot data a composite score was computed by combining the scores for the binary- and ternary-relational items. The mean composite score for the gambling, future-oriented and conditional discrimination/reversal learning (based on the number of items correct on problem 1) tasks are presented in Table 9.3. Zero-order correlations are shown below the diagonal, and age-partialled correlations are shown above the diagonal. The zero-order correlations indicated significant associations between the three variables, and each was also significantly correlated with age. Only the correlation between the gambling task and future-oriented decision-making remained significant when age in months was partialled out.
Table 9.3

Zero-order (below diagonal) and Age-partialled (above diagonal) Correlations Among the Three Hot Executive Measures in the Pilot Study (N = 72)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gambling</td>
<td></td>
<td>.24*</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>2. Future-Oriented</td>
<td>.44***</td>
<td></td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td>3. CD/RL</td>
<td>.39**</td>
<td>.55***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Age (months)</td>
<td>.45***</td>
<td>.60***</td>
<td>.72***</td>
<td></td>
</tr>
</tbody>
</table>

Mean

|      | 55.68 | 14.90 | 34.85 | 53.94 |

Standard Deviation

|      | 13.20 | 4.37  | 7.21  | 8.88  |

Note. *p < .05 **p = .001. ***p < .001

Main Study

Composite scores were computed by combining the scores for the binary- and ternary-relational items of each of the four hot and three cool executive functions tasks. The mean composite score for each executive task, zero-order correlations and age-partialled correlations are presented in Table 9.4. Significant correlations were found between all measures of hot executive function (ranging from .50 to .80). Significant associations were also found within the cool domain (ranging from .68 to .73). The correlations between tasks across the domains were as strong as those within the two domains (ranging from .50 to .78). The high correlations among the tasks support the existence of a domain general ability that underlies performance in both hot and cool executive function tasks, that is, the ability to process complex relations. Table 9.4 demonstrates that 2 out of 6 correlations within the hot domain, 2 out of 3 correlations within the cool domain, and 4 out of 12 across domain correlations remained significant after age was partialled out.
### Table 9.4

*Mean Composite Score, Zero-order and Age-partialled Correlations Among the Executive Measures and Age (N = 120)*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Gambling</td>
<td>.24*</td>
<td>.03</td>
<td>.14</td>
<td>.13</td>
<td>.12</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Future-Oriented</td>
<td>.59***</td>
<td>.06</td>
<td>.25**</td>
<td>.10</td>
<td>.15</td>
<td>.31***</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>CD/RL</td>
<td>.50***</td>
<td>.61***</td>
<td>.39***</td>
<td>.16</td>
<td>.33***</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>TOM</td>
<td>.59***</td>
<td>.72***</td>
<td>.80***</td>
<td>.17</td>
<td>.27**</td>
<td>.22*</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Transitivity</td>
<td>.59***</td>
<td>.67***</td>
<td>.73***</td>
<td>.78***</td>
<td>.10</td>
<td>.21*</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Class Inclusion</td>
<td>.56***</td>
<td>.66***</td>
<td>.76***</td>
<td>.78***</td>
<td>.73***</td>
<td></td>
<td>.24**</td>
</tr>
<tr>
<td>7.</td>
<td>DCCS</td>
<td>.50***</td>
<td>.68***</td>
<td>.60***</td>
<td>.69***</td>
<td>.70***</td>
<td>.68***</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Age (months)</td>
<td>.62***</td>
<td>.74***</td>
<td>.79***</td>
<td>.86***</td>
<td>.86***</td>
<td>.81***</td>
<td>.72***</td>
</tr>
</tbody>
</table>

Mean: 95.67 16.65 78.38 11.14 9.72 7.49 14.53 59.62

Standard Deviation: 14.92 4.52 12.11 2.91 3.12 3.60 1.72 13.46

Note.  *p < .05.  **p < .01.  ***p < .001.
The significant correlations found between the cool executive function tasks is consistent with the results of Andrews and Halford (2002) who reported positive correlations between transitivity and class inclusion, and Hongwanishkul et al (2005) who found a positive correlation between the DCCS and a self-ordered pointing task. Similarly, the correlations between the hot tasks are consistent with Moore et al. (1998) who found significant correlations between theory-of-mind performance and measures of future-oriented altruism. Cross domain associations have been reported previously. Across three experiments, Andrews et al., (2003) found significant correlations between theory-of-mind performance and transitivity, DCCS and class inclusion.

In contrast to the current research, Hongwanishkul et al (2005) found a negative correlation between two hot tasks, the Children’s Gambling task and a delay of gratification task. The delay task involved three types of rewards (stickers, candies, and pennies) and three types of choices (1 now versus 2 later, 1 now versus 4 later, and 1 now versus 6 later), which was administered to 3- to 5-year-old children. The results indicated a negative association (-.22 when age was partialled) between this version of the task and the Children’s Gambling Task. In contrast, both the pilot and main study of the current research found a significant positive correlation between two comparable measures (.21 and .24 respectively with age-partialled). The differences across the findings of the two studies may be attributable to the slightly different task designs and dependent variables used. However, it again raises questions regarding the 'hotness' of the future-oriented decision-making task.

In order to further investigate a possible differential age of emergence for hot versus cool executive functions, the tasks from the individual domains were subjected to a data reduction technique to allow for a comparison of individual factor scores. The four hot tasks were first subjected to a factor analysis with principal axis extraction.
This technique was chosen as it is based on common variance, and should reflect what is common to the tasks whose surface characteristics differ. Subsequent analyses indicated that the choice of extraction method did not alter the results.

Prior to performing the analysis, the suitability of the data for factor analysis was assessed. Inspection of the correlation matrix revealed all coefficients were .3 and above. The Kaiser-Meyer-Olkin value was .93, exceeding the recommended value of .6 and the Bartlett’s Test of Sphericity ($p < .001$) supported the factorability of the correlation matrix. For the hot executive function tasks, all four variables loaded on a single factor (eigenvalue = 2.91). This component was labeled Hot EF and it explained 64.60% of the variance. The factor loadings are shown in Table 9.5.

Table 9.5

*Loadings of the Hot Executive Function Factor*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hot EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory-of-mind</td>
<td>.94</td>
</tr>
<tr>
<td>Conditional Discrimination</td>
<td>.81</td>
</tr>
<tr>
<td>Future-Oriented</td>
<td>.80</td>
</tr>
<tr>
<td>Gambling</td>
<td>.66</td>
</tr>
</tbody>
</table>

A second principal axis analysis was conducted on the three cool measures. Inspection of the correlation matrix revealed all coefficients were .3 and above. The Kaiser-Meyer-Olkin value of .89, and Bartlett’s Test of Sphericity ($p < .001$), supported the factorability of the correlation matrix. For the cool executive function domain, all three tasks loaded on a single factor (eigenvalue = 2.41). This component was labelled
Cool EF and explained 70.47% of the variance. The factor loadings are presented in Table 9.6.

Table 9.6

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cool EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitivity</td>
<td>.86</td>
</tr>
<tr>
<td>Class Inclusion</td>
<td>.85</td>
</tr>
<tr>
<td>DCCS</td>
<td>.81</td>
</tr>
</tbody>
</table>

For each of the preceding analyses, factor scores were computed using the regression approach. The factor scores were subjected to a 4 (age: 3, 4, 5, 6) × 2 (domain: hot, cool) ANOVA. The results indicated a significant main effect of age $F(3, 116) = 246.90, p < .001, \eta^2 = .865$, and a significant Age × Domain interaction, $F(3, 116) = 4.64, p = .004, \eta^2 = .107$. Means and standard errors for the interaction are presented in Figure 9.1. Paired samples $t$-tests indicated significant differences between the two domains for 4-year-olds, $t(29) = 2.62, p = .014$. These children performed better in the hot domain than the cool domain. There were no significant differences between the two domains for 3-year-olds, $t(29) = -1.58, p = .124$, 5-year-olds, $t(29) = .66, p = .511$ or 6-year-olds, $t(29) = -1.91, p = .067$. 
Figure 9.1. Age effects on the hot and cool domain factor scores.

Conclusions Regarding the Emergence of Hot and Cool Executive Functions

The results of the above analyses reveal consistent development of executive abilities during the preschool years. However, differences may be noted between the development of hot versus cool executive functions. The patterns of age-related improvement, pass-fail scores and age effects on the domain factor scores all suggest that while hot executive functions may begin to develop around 4-years of age, similar levels of improvement are not seen in cool executive functions until 5-years of age. Thus, success on the ternary-relational items of hot executive function tasks appears to emerge slightly earlier than success on items at the same level of complexity in the cool executive function domain.
CHAPTER 10. CONCLUSIONS

The current research examined performance in two theorised domains of executive function, using developmentally appropriate tasks which were equivalent in complexity. Consistent with the developmental trends observed for the frontal lobes during the ages of three to six years, performance on both the hot and cool executive functions improved significantly across these ages. The implications of these results for the applicability of complexity theories to the assessment of executive functions associated with the OFC and the DL-PFC, and the implications for the hot-cool distinction, will be considered.

**Theoretical Implications**

*Complexity Theories*

*Relational Complexity.* The findings of the current research are consistent with relational complexity theory in three main ways. Firstly, significant complexity effects were found across all seven tasks. Items at a higher level of complexity were experienced as relatively more difficult by the majority of children, while other factors were tightly controlled. The significant complexity effects observed within each task provide support for the underlying relational complexity metric and the complexity analyses proposed in the current study.

Secondly, the age of emergence of the ability to pass the binary- and ternary-relational items of each task is consistent with predictions of relational complexity theory. Performance on all tasks improved with age, but the age effects were strongest on the ternary-relational items. The pass-fail data indicated that the majority of children in all age groups succeeded on the binary-relational items. However, it was not until a median of five years of age that children were able to process ternary relations.
Consequently, the ternary-relational items produce the greatest differences in performance between the four age groups.

Thirdly, the relational complexity metric appears to apply to both hot and cool executive functions, with items of higher relational complexity experienced as more difficult than items of lower relational complexity in tasks associated with both domains. The success of 3-year-olds on the binary-relational items of all tasks indicates that the difficulty experienced by the younger children on the ternary-relational items, is not due to task procedures since these were comparable for both levels of complexity. Rather, it appears that the complexity of the particular tasks was a major source of difficulty. Thus, complexity is an important factor underlying children’s performance on hot and cool tasks and should be taken into account when assessing the development of executive functions.

**CCC Theory.** The performance of children in the current study might also be able to be explained in terms of the CCC theory. As described in the previous chapters, both the Children’s Gambling Task and the DCCS have been analysed in terms of the ability to integrate higher order rules that allow for the deliberate selection of lower order rules (e.g., Kerr & Zelazo, 2004; Zelazo & Frye, 1997). Similarly, the background variable involved in the conditional discrimination task can be interpreted as a setting condition that allows for two incompatible pairs of rules to be integrated into a single rule system. The hierarchical structure of the class inclusion task also allows for the possibility of an interpretation based on the CCC theory. An interpretation of transitive inference is less obvious, however, may be possible if the transitivity task is conceptualised in term of a higher order rule, for example “monotonically higher than”. It remains to be seen, however, whether the CCC approach could also be applied to the future-oriented decision-making task.
Development of Executive Functions

**Hot-Cool Distinction**

As described in Chapter 9, success on the ternary-relational items of the hot executive function tasks appears to emerge earlier than success on the ternary-relational items of the cool executive function tasks. This is consistent with the earlier development of the OFC than the DL-PFC. However, the high correlations between the hot and cool executive function tasks support the existence of a domain general ability that underlies performance of all tasks. As complexity was manipulated across the seven tasks, this appears to be the ability to process complex relations.

Although the hot-cool distinction differentiates between processes associated with the DL-PFC and OFC, Zelazo, Qu and Müller (2004) point out that they remain part of a single coordinated system and most often work together. Additionally, they suggested that a possible strategy for solving some hot problems might be to reconceptualise them in decontextualised terms in order to address them using cool executive functions. Thus while individual tasks might emphasise either hot or cool executive functions, it might be impossible to design a task that is a pure measure of one domain only. If hot and cool tasks are to be conceptualized as part of a single executive system, perhaps at different ends of a hot/cool continuum, the strong intercorrelations found in the current results is to be expected.

An alternative explanation for the earlier emergence of executive functions associated with the OFC, is that the hot tasks may inherently be more interesting to young children. Three tasks from the hot domain involved rewards including chocolates (Children’s Gambling Task), stickers (Future-oriented decisions) or positive feedback (receiving smiley faces). The fourth task, theory-of-mind, often involved stories being acted out or the use of play materials. While these methods were inherent in the design of tasks tapping the OFC, such task procedures and rewards may be more motivating and foster greater attention, than the cool executive function tasks, particularly in
younger children. Several attempts were made by the experimenter to maintain children’s motivation across all testing items. However, the possibility remains that the superior performance of 4-year-olds on the hot tasks (relative to the cool tasks) reflects greater attention and motivation rather than the involvement of the OFC. Future research could examine whether performance of cool executive function tasks could be improved by introducing a motivational component into their design.

**Methodological Issues**

One difficulty encountered during the research was the lack of a controlled testing environment. Children were tested on the premises of their day-care or schools as this was an efficient way to obtain access to the children, and it was also an environment in which they were comfortable. However, the testing sessions were occasionally interrupted by staff members and curious children. The distractions were random and sporadic, and as such there is no reason to suspect that the results were influenced in a systematic way. A further difficulty was produced by the number of tasks administered in the current study. Although testing was divided into as many sessions as were required by the individual children, the attention of the younger children began to wane towards the end of each session. Due to the counterbalanced order of the tasks, the drop in attention should not have had any systematic effect on the results. The high level of performance across the binary-relational items of the seven tasks also indicates that the performance of children did not appear to be affected in any systematic way. However, a reduction in the number of tasks, or items within tasks, might be suggested for further investigations.

Although strong support was found for the applicability of relational complexity theory to both hot and cool domains of executive function, several recommendations can be made to allow for further refinement of two of the hot tasks that were adapted in the current study.
Gambling Task. A potential alternative interpretation of the gambling task results is the variability of the losses. While the original ternary-relational version included four possible loss scenarios (i.e., 0, 4, 5, 6), the binary-relational versions included only two possible loss outcomes (0 or 5). Thus, the complexity manipulation was confounded with the variability of losses. The limited variability of losses in the binary-relational versions may have aided children in identifying the advantageous deck. This interpretation could be tested by equating the variability of losses across both levels of complexity, for example by using losses of 0 or 5 for both the binary- and ternary-relational items.

Future-oriented decision-making. As described in the previous chapter, the patterns of age-related improvement and pass/fail percentages of the ternary-relational items produced a mixed picture regarding the hot-cool status of the future-oriented decision-making task. Metcalfe and Mischel (1999) suggested that episodic representation and thought is primarily a function of the cool, cognitive know system. The items utilised in the future-oriented task could be seen as requiring the ability to project into the future. Therefore, the involvement of episodic representation in the performance of this task may cast doubt over its ‘hotness’ and associated involvement of the OFC.

As aforementioned, the binary-relational items of the future-oriented decision making task did not entail a delay variable. While the high correlations between the future-oriented decision-making tasks and the other executive function tasks indicate that the complexity manipulation in these items worked as intended, and in the same manner as in the other tasks, future research would be beneficial. Incorporation of a delay variable into binary-relational items would provide further support for the notion that young children can understand all of the individual components of the ternary-relational items, however, experience difficulty integrating the components.
While executive function deficits have consistently been found in a range of developmental disorders, it has been suggested that the severity and profile of such deficits may be distinct for each disorder. Specific comparisons have been made between the executive profiles of children with autism versus children with ADHD. More severe executive function deficits relative to IQ matched controls are found in autism than in ADHD on global executive function tasks (Pennington & Ozonoff, 1996). In terms of profile, Zelazo and Müller’s (2002) hot-cool distinction argues that important aspects of autism may be understood in terms of a primary deficit in the affective ‘hot’ aspects of executive function, which leads to associated impairments in cool executive functions. On the other hand, important aspects of ADHD may be understood in terms of a primary deficit in the cognitive ‘cool’ aspects, while the ‘hot’ aspects are relatively unimpaired.

To date, research regarding the primacy of executive function impairments in these disorders has been limited as it has not taken complexity into account. That is, tasks currently used in investigations of executive function profiles might differ in terms of the complexity of the cognitive processes that the tasks require (Zelazo & Müller, 2002). Therefore comparisons across tasks might be misleading because these tasks vary in terms of the demands they place on participants. This difficulty could be addressed by administering the current test battery of hot and cool executive function tasks (each having complexity controlled) to a range of clinical samples including children with autism and ADHD. This would provide the opportunity to not only further examine the development of hot/cool executive functions and the role of relational complexity in the performance of these tasks, but the design would allow the hot-cool distinction and complexity level to be decoupled. That is, if different clinical populations are individually impaired in terms of hot or cool deficits, they should be
impaired on binary-relational items as well as ternary-relational items in that domain. However, if a clinical population is impaired on complexity, they might succeed on binary-relational items in both hot and cool domains but fail on ternary-relational items.

The future administration of the hot-cool battery to clinical samples would help to clarify the discriminant validity problem associated with the impairments of executive functions across multiple disorders. The current research, however, has also contributed to the understanding of executive functions more broadly. Complexity appears to be a critical factor underlying children’s performance on both hot and cool executive function tasks. Future assessment regarding the development of executive abilities might also benefit from recognising the contribution of the relational complexity metric.
REFERENCES


D’Esposito, M. D., Aguirre, G. K., Zarahn, E., Ballard, D., Shin, R. K., & Lease, J.
*Cognitive Brain Research, 7*, 1-13.

York: Puttman’s Sons.

*Journal of Verbal Learning and Verbal Behaviour, 19*, 450-466.


*Neuropsychologia, 33*, 1243-1253.

analysis and modelling within a neural network. *Cerebral Cortex, 1*, 62-79.

San Antonio, TX: The Psychological Corporation.


with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia, 38*, 388-402.


APPENDICES

Appendix A: Information Letters and Informed Consent

INFORMATION SHEET FOR PARENTS/CAREGIVERS

A Relational Complexity Approach to the Development of Hot/Cool Executive Functions in Children

Investigators:

Katie Bunch, School of Applied Psychology, Griffith University, Gold Coast.
Phone: 07 5594 9717; Email: k.bunch@griffith.edu.au
Dr. Glenda Andrews, School of Applied Psychology, Griffith University, Gold Coast.
Phone: 07 5552 8613; Email: g.andrews@griffith.edu.au

This research is being conducted by Katie Bunch, as part of her Clinical PhD course requirements. The project is being supervised by Dr. Glenda Andrews from the School of Applied Psychology at the Gold Coast campus at Griffith University. This study is designed to assess the development of children’s thinking, in particular, their ability to plan, remember and execute actions and how this develops between the ages of 3- and 7-years of age. Children will be required to complete a number of tasks, presented in game formats that most children enjoy. These include “games” in which children will be asked to sort cards, distribute stickers, identify shapes and colours, and answer questions regarding the numerosity of groups of objects. The tasks will take approximately 2 hours to complete. The tasks will be administered in two sessions, however this may be altered according to the needs of each individual child. The research will be conducted at the daycare, kindergarten center or school during normal class hours. The project is research and not part of the curriculum or normal school activities.

Some of the tasks involve small rewards being given to your child. These rewards include mini M&M chocolates, or stickers to be placed in a sticker book. You may choose the type of reward to be given to your child by ticking the appropriate box on the consent form. If you do not wish your child to receive chocolate as a reward, you should tick the ‘stickers only’ option.

There are no risks to your child associated with participation in this study. You may withdraw your child from the study at any time, and testing will cease immediately if your child indicates they do not wish to continue. The confidentiality of your child’s responses is assured as data from each participant will be identifiable only by a participant number. The data for all the participants will be combined and no data about individual participants will be reported in any publications or presentations of the results of this research. If you need further clarification of any of the issues addressed above, or if you have any additional questions, please contact Katie Bunch (contact details are given above).

If you have any complaints concerning the manner in which the research project is conducted, you may contact the Manager, Research Ethics, Office for Research, Bray Centre, Griffith University, Kessels Road, Nathan, Qld 4111, telephone (07) 38755585 or email research-ethics@griffith.edu.au.

A summary of the overall outcomes of the research will be made available on completion of the work. Participants may obtain this from the investigator. Katie Bunch, Dr Glenda Andrews and Griffith University would like to thank you for taking time to read this request. It would assist our research greatly if you would allow you child to participate.

If you agree to your child’s participation, please complete the consent form and return it to your child’s daycare, kindergarten or school at your next opportunity.
A Relational Complexity Approach to the Development of Hot/Cool Executive Functions in Children

I agree to let my child participate in the *A relational complexity approach to the development of hot/cool executive functions in children* project and give my consent freely. I understand that my child is not required to participate in this research project if I do not wish it and that I can withdraw him/her from the study at any time without needing to explain my reasons for withdrawing. No loss of benefit or treatment will occur as a result of my withdrawal of him/her and no penalty will be incurred.

I have read the information sheet and consent form. I understand that my child will be asked to complete a number of tasks, which are presented in game formats by a trained person, in several sessions at the childcare center/school during normal class hours. I understand that small rewards may be given to my child, and that I will choose the type of reward that they will receive. I understand that my child’s responses will remain confidential and that data for all participants will be combined prior to being reported in any publications. I understand that a summary of the overall outcomes of the research will be available to me on completion of the work.

I understand that the study will be carried out as described in the information statement, a copy of which I have retained. I have had all questions answered to my satisfaction.

Signatures

Investigator: ........................................ Date: ........................................

Parent/Caregiver: ........................................ Date: ........................................

Child’s Name: ........................................ Child’s Date-of-birth: ........................................

*Please check appropriate box:*

☐ My child is to be rewarded with both a small amount of mini M&Ms and stickers.

☐ My child is to be rewarded with stickers only.
## Appendix B: Response Sheet for the Gambling Task

### Identification:
- Date of birth: 
- Age: 
- Order: 3

### Ternary Level

| Deck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | Best |
|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|
| Adv. |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | Best |
| Disadv. |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |

**Total Advantageous Choices =**

### Binary Level Pack A

| Deck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | Best |
|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|
| Adv. |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | Best |
| Disadv. |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |

**Total Advantageous Choices =**

### Binary Level Pack B

| Deck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | Best |
|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|
| Adv. |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | Best |
| Disadv. |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |

**Total Advantageous Choices =**
## Appendix C: Response Sheet for Future-Oriented Decision-Making Task

**Participant ID:** _____  **DOB:** _____  **Order:** 1  **Age:** _____

### Binary level items

<table>
<thead>
<tr>
<th>#</th>
<th>ITEM TYPE</th>
<th>OPTION 1</th>
<th>OPTION 2</th>
<th>RESPONSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Shared gratification</td>
<td>1 each now</td>
<td>1 for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>E1</td>
<td>Self gratification</td>
<td>1 for self now</td>
<td>2 for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>O1</td>
<td>Other gratification</td>
<td>2 for other now</td>
<td>1 for other now</td>
<td>0 1</td>
</tr>
<tr>
<td>E2</td>
<td>Self gratification</td>
<td>4 for self now</td>
<td>2 for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>O2</td>
<td>Other gratification</td>
<td>2 for other now</td>
<td>4 for other now</td>
<td>0 1</td>
</tr>
<tr>
<td>S2</td>
<td>Shared gratification</td>
<td>2 for self now</td>
<td>2 each now</td>
<td>0 1</td>
</tr>
<tr>
<td>O3</td>
<td>Other gratification</td>
<td>2 for other now</td>
<td>1 for other now</td>
<td>0 1</td>
</tr>
<tr>
<td>S3</td>
<td>Shared gratification</td>
<td>3 for self now</td>
<td>3 each now</td>
<td>0 1</td>
</tr>
<tr>
<td>E3</td>
<td>Self gratification</td>
<td>2 for self now</td>
<td>1 for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>S4</td>
<td>Shared gratification</td>
<td>4 each now</td>
<td>4 for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>E4</td>
<td>Self gratification</td>
<td>2 for self now</td>
<td>4 for self now</td>
<td>0 1</td>
</tr>
</tbody>
</table>

### Ternary level items

<table>
<thead>
<tr>
<th>#</th>
<th>ITEM TYPE</th>
<th>OPTION 1</th>
<th>OPTION 2</th>
<th>RESPONSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Shared gratification</td>
<td>1 sticker each later</td>
<td>1 sticker for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>E1</td>
<td>Self gratification</td>
<td>2 stickers for self later</td>
<td>1 sticker for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>O1</td>
<td>Other gratification</td>
<td>1 sticker for other now</td>
<td>2 stickers for other later</td>
<td>0 1</td>
</tr>
<tr>
<td>E2</td>
<td>Self gratification</td>
<td>4 stickers for self later</td>
<td>2 stickers for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>O2</td>
<td>Other gratification</td>
<td>4 stickers for other later</td>
<td>2 stickers for other now</td>
<td>0 1</td>
</tr>
<tr>
<td>S2</td>
<td>Shared gratification</td>
<td>2 stickers for self now</td>
<td>2 stickers each later</td>
<td>0 1</td>
</tr>
<tr>
<td>O3</td>
<td>Other gratification</td>
<td>2 stickers for other later</td>
<td>1 stickers for other now</td>
<td>0 1</td>
</tr>
<tr>
<td>S3</td>
<td>Shared gratification</td>
<td>3 stickers for self now</td>
<td>3 stickers each later</td>
<td>0 1</td>
</tr>
<tr>
<td>E3</td>
<td>Self gratification</td>
<td>1 stickers for self now</td>
<td>2 stickers for self later</td>
<td>0 1</td>
</tr>
<tr>
<td>S4</td>
<td>Shared gratification</td>
<td>4 stickers each later</td>
<td>4 stickers for self now</td>
<td>0 1</td>
</tr>
<tr>
<td>E4</td>
<td>Self gratification</td>
<td>2 stickers for self now</td>
<td>4 stickers for self later</td>
<td>0 1</td>
</tr>
<tr>
<td>O4</td>
<td>Other gratification</td>
<td>2 stickers for other now</td>
<td>4 stickers for other later</td>
<td>0 1</td>
</tr>
</tbody>
</table>
### Appendix D: Response Sheet for the Transformation and Connections Tasks

<table>
<thead>
<tr>
<th>Transformation Tasks</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Change of location</strong></td>
<td>Where would the dog be if that happened, A or B?</td>
<td>A</td>
<td>B</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>Where would the bear be if that happened, C or A?</td>
<td>A</td>
<td>C</td>
<td>0 1</td>
</tr>
<tr>
<td><strong>2. Colour change of appearance.</strong></td>
<td>What colour do things look through this plastic, yellow or red?</td>
<td>YELLOW</td>
<td>RED</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>What colour do things look through this plastic, purple or green?</td>
<td>GREEN</td>
<td>PURPLE</td>
<td>0 1</td>
</tr>
<tr>
<td><strong>3. Appearance – wobbly</strong></td>
<td>How do things look through this water, still or wobbly?</td>
<td>STILL</td>
<td>WOBBLY</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>How do things look through this water, wobbly or still?</td>
<td>STILL</td>
<td>WOBBLY</td>
<td>0 1</td>
</tr>
<tr>
<td><strong>4. Change of contents</strong></td>
<td>What would be in the box if that happened, biscuits or paper?</td>
<td>BISCUITS</td>
<td>PAPER</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>What would be in the box if that happened, stones or biscuits?</td>
<td>STONES</td>
<td>BISCUITS</td>
<td>0 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Connections Tasks</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Open-ended box</strong></td>
<td>Look, can you see what is inside the box right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td>(a) Open end face R</td>
<td>How about me. Can I see what is inside the box right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td>(b) Open end facing child</td>
<td>Look, can you see what is inside the box right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>How about me. Can I see what is inside the box right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td><strong>2. Drum Item</strong></td>
<td>Look, do you see the drum right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td>(a) Child’s side</td>
<td>How about me. Do I see the drum right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td>(b) Researcher’s side</td>
<td>Look, do you see the drum right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>What about me, Do I see the drum right now?</td>
<td>YES</td>
<td>NO</td>
<td>0 1</td>
</tr>
</tbody>
</table>
Appendix E: Response Sheet for Appearance-Reality and False-Belief Tasks

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>DOB:</th>
<th>Order</th>
<th>Age:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

**Appearance Reality Tasks**

1. **Fish Filter**
   - A: When you look at this fish right now, does the fish look white or does it look blue?  
     - **WHITE** | **BLUE** | 0 1
   - R: What colour if the fish really and truly, blue or white?  
     - **WHITE** | **BLUE** | 0 1

2. **Skewer in Glass**
   - R: What is this stick really and truly? Is it straight or is it bent and cooked?  
     - **BENT** | **STRAIGHT** | 0 1
   - A: When you look at this with your eyes right now, does this stick look bent and crooked or does it look straight?  
     - **BENT** | **STRAIGHT** | 0 1

3. **Crayon Box**
   - A: When you look at this box with your eyes right now, what does this box look like it has in it, sticks or crayons?  
     - **STICKS** | **CRAYONS** | 0 1
   - R: What’s really and truly in this box, crayons or sticks?  
     - **STICKS** | **CRAYONS** | 0 1

**False Belief Tasks**

1. **Smarties Box**
   - R: What is really and truly in the box, pencils or smarties?  
     - **SMARTIES** | **PENCILS** | 0 1
   - B: If someone came into this room right now and had not seen what was in the box, what would they think was in it, smarties or pencils?  
     - **SMARTIES** | **PENCILS** | 0 1

2. **Colour Cat**
   - B: When you first saw the cat, all covered up like this, what colour did you think the cat really was, green or white?  
     - **GREEN** | **WHITE** | 0 1
   - R: What colour did you think it was really and truly, white or green?  
     - **GREEN** | **WHITE** | 0 1

3. **Chocolate Hiding Task**
   - R: Where is the chocolate really and truly, in the green or the blue cupboard?  
     - **BLUE** | **GREEN** | 0 1
   - B: Where will Maxi look first for the chocolate, in the blue cupboard or in the green cupboard?  
     - **BLUE** | **GREEN** | 0 1
Appendix F: Response Sheet for Transitivity

Participant identification:
Age:
Date of birth:
Date of test:
School:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P1B</td>
<td>Green (B) Purple(C)</td>
<td>Yellow(D)</td>
<td>Blue(E)</td>
<td>Red(A)</td>
<td>Circle response.</td>
</tr>
<tr>
<td>P2B</td>
<td>Purple (C) Blue (D)</td>
<td>Yellow(B)</td>
<td>Red(A)</td>
<td>Green(E)</td>
<td>Blue Yellow</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P1T</td>
<td>Purple (B) Red (D)</td>
<td>Blue (C)</td>
<td>Green (E)</td>
<td>None/Step 2/Step 3</td>
<td></td>
</tr>
<tr>
<td>P2T</td>
<td>Yellow (D) Blue (B)</td>
<td>Green (A)</td>
<td>Red (C)</td>
<td>None/Step 2/Step 3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Test1B</td>
<td>Purple (C) Green (D)</td>
<td>Red (B)</td>
<td>Yellow(E)</td>
<td>Blue (A)</td>
<td>Circle response</td>
</tr>
<tr>
<td>Test2B</td>
<td>Green (C) Red (B)</td>
<td>Yellow (D)</td>
<td>Purple (A)</td>
<td>Blue (E)</td>
<td>Yellow?Red</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Test1T</td>
<td>Blue(B) ? Purple(D)</td>
<td>Yellow (C)</td>
<td>Green (A)</td>
<td>BD + C</td>
</tr>
<tr>
<td>Test2T</td>
<td>Purple(D) ? Green(B)</td>
<td>Yellow (E)</td>
<td>Blue (C)</td>
<td></td>
</tr>
<tr>
<td>Test3T</td>
<td>Red(B) ? Yellow (D)</td>
<td>Purple (A)</td>
<td>Green (C)</td>
<td></td>
</tr>
<tr>
<td>Test4T</td>
<td>Red(D) ? Green(B)</td>
<td>Blue (C)</td>
<td>Purple (E)</td>
<td></td>
</tr>
<tr>
<td>Test5T</td>
<td>Purple(B) ? Yellow(D)</td>
<td>Red (C)</td>
<td>Green (A)</td>
<td></td>
</tr>
<tr>
<td>Test6T</td>
<td>Green (D) ? Red (B)</td>
<td>Blue (A)</td>
<td>Purple (C)</td>
<td></td>
</tr>
<tr>
<td>Test7T</td>
<td>Blue (B) ? Red (D)</td>
<td>Purple (E)</td>
<td>Green (C)</td>
<td></td>
</tr>
<tr>
<td>Test8T</td>
<td>Blue(D) ? Yellow (B)</td>
<td>Purple (C)</td>
<td>Red (A)</td>
<td></td>
</tr>
</tbody>
</table>

**Binary score** out of 8 (Step1 + Step 2 + Step 3 + Step 4) =

**Ternary score** out of 8 (# items where BD and C were correct) =

Comments:
## Appendix G: Response Sheet for Class Inclusion

<table>
<thead>
<tr>
<th>Identification</th>
<th>Date of birth</th>
<th>Order: Forward</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CI1 Identify</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Green things</td>
<td>Yellow things</td>
</tr>
<tr>
<td>Are all the green things circles?</td>
<td>Yes/No</td>
<td></td>
</tr>
<tr>
<td>Are all the yellow things circles?</td>
<td>Yes/No</td>
<td></td>
</tr>
<tr>
<td><strong>Are there more?</strong></td>
<td>(Circle one)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Yellow things</td>
<td>Green things</td>
</tr>
<tr>
<td>C</td>
<td>Circles</td>
<td>Green things</td>
</tr>
<tr>
<td>B</td>
<td>Yellow things</td>
<td>Circles</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **CI2 Identify** |               |               |
|                  | Red things    | Blue things   | Triangles |
| Are all the red things triangles? | Yes/No |
| Are all the blue things triangles? | Yes/No |
| **Are there more?** | (Circle one) |               |
| A               | Red things    | Blue things   |         |
| B               | Red things    | Triangles     |         |
| C               | Blue things   | Triangles     |         |
| Totals          |               |               |

| **CI3 Identify** |               |               |
|                  | Yellow things | Blue things   | Squares |
| Are all the yellow things squares? | Yes/No |
| Are all the blue things squares? | Yes/No |
| **Are there more?** | (Circle one) |               |
| A               | Yellow things | Blue things   |         |
| C               | Yellow things | Squares       |         |
| B               | Squares      | Blue things   |         |
| Totals          |               |               |

| **CI4 Identify** |               |               |
|                  | Circle        | Square        | Green things |
| Are all the circles green things? | Yes/No |
| Are all the squares green things? | Yes/No |
| **Are there more?** | (Circle one) |               |
| A               | Squares      | Circle        |         |
| B               | Circles     | Green things  |         |
| C               | Green things | Squares       |         |
| Totals          |               |               |

| **CI5 Identify** |               |               |
|                  | Triangles     | Squares       | Red things |
| Are all the triangles red things? | Yes/No |
| Are all the squares red things? | Yes/No |
| **Are there more?** | (Circle one) |               |
| A               | Triangles     | Squares       |         |
| B               | Red things    | Squares       |         |
| C               | Triangles     | Red things    |         |
| Totals          |               |               |

| **CI6 Identify** |               |               |
|                  | Triangles     | Circles       | Blue things |
| Are all the triangles blue things? | Yes/No |
| Are all the circles blue things? | Yes/No |
| **Are there more?** | (Circle one) |               |
| A               | Circles      | Triangles     |         |
| C               | Blue things  | Circles       |         |
| B               | Blue things  | Triangles     |         |
| Totals          |               |               |
### Appendix H: Standards for Templates and Objects Sorted in the Binary and Ternary-Relational DCCS

<table>
<thead>
<tr>
<th>Template</th>
<th>Standard</th>
<th>Position</th>
<th>Standard</th>
<th>Position</th>
<th>Objects Sorted Ternary</th>
<th>Objects Sorted Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>green triangle</td>
<td>left column</td>
<td>red circle</td>
<td>right column</td>
<td>red triangle, green circle</td>
<td>red circles, green circle</td>
</tr>
<tr>
<td>1b</td>
<td>green triangle</td>
<td>right column</td>
<td>red circle</td>
<td>left column</td>
<td>red triangle, green circle</td>
<td>red triangle, green triangle</td>
</tr>
<tr>
<td>2a</td>
<td>red triangle</td>
<td>top row</td>
<td>green circle</td>
<td>bottom row</td>
<td>green triangle, red circle</td>
<td>red triangle, red circle</td>
</tr>
<tr>
<td>2b</td>
<td>red triangle</td>
<td>bottom row</td>
<td>green circle</td>
<td>top row</td>
<td>green triangle, red circle</td>
<td>green triangle, green circle</td>
</tr>
<tr>
<td>3a</td>
<td>blue flower</td>
<td>left column</td>
<td>yellow boat</td>
<td>right column</td>
<td>yellow flower, blue boat</td>
<td>yellow boat, yellow flower</td>
</tr>
<tr>
<td>3b</td>
<td>blue flower</td>
<td>right column</td>
<td>yellow boat</td>
<td>left column</td>
<td>yellow flower, blue boat</td>
<td>blue boat, blue flower</td>
</tr>
<tr>
<td>4a</td>
<td>yellow flower</td>
<td>top row</td>
<td>blue boat</td>
<td>bottom row</td>
<td>blue flower, yellow boat</td>
<td>yellow flower, blue flower</td>
</tr>
<tr>
<td>4b</td>
<td>yellow flower</td>
<td>bottom row</td>
<td>blue boat</td>
<td>top row</td>
<td>blue flower, yellow boat</td>
<td>yellow boat, blue boat</td>
</tr>
</tbody>
</table>
## Appendix I: Response Sheets for the DCCS

### Card-Sorting Task - Ternary Level

<table>
<thead>
<tr>
<th>Display</th>
<th>Stimuli</th>
<th>Instructions</th>
<th>Child's Response</th>
<th>Order</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp. 1a</td>
<td>Demo</td>
<td>Red Triangle</td>
<td>See, I have a red one. It goes here because this one is red too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Green Circle</td>
<td>See, I have a green one. It goes here because this one is green too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 1b</td>
<td>Test</td>
<td>Red Triangle</td>
<td>Here is a red one, put it in the space where it belongs.</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Green Circle</td>
<td>Here is a green one, put it in a space where it belongs.</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 2a</td>
<td>Demo</td>
<td>Green Triangle</td>
<td>See, I have a triangle. It goes here because this one is a triangle too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>Red Circle</td>
<td>See, I have a circle. It goes here because this one is a circle too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 2b</td>
<td>Test</td>
<td>Green Triangle</td>
<td>Here is a triangle, put it in the space where it belongs.</td>
<td>Top</td>
<td>Bottom</td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>Red Circle</td>
<td>Here is a circle, put it in the space where it belongs.</td>
<td>Top</td>
<td>Bottom</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 3b</td>
<td>Test</td>
<td>Blue boat</td>
<td>Here is a boat, put it in the space where it belongs.</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>Yellow flower</td>
<td>Here is a flower, put it in the space where it belongs.</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 3a</td>
<td>Demo</td>
<td>Blue boat</td>
<td>See, I have a boat. It goes here because this one is a boat too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>Yellow flower</td>
<td>See, I have a flower. It goes here because this one is a flower too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 3b</td>
<td>Re-test</td>
<td>Blue boat</td>
<td>Here is a boat, put it in the space where it belongs.</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>Yellow flower</td>
<td>Here is a flower, put it in the space where it belongs.</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 4b</td>
<td>Test</td>
<td>Yellow boat</td>
<td>Here is a yellow thing, put it in the space where it belongs.</td>
<td>Top</td>
<td>Bottom</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Blue flower</td>
<td>Here is a blue thing, put it in the space where it belongs.</td>
<td>Top</td>
<td>Bottom</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 4a</td>
<td>Demo</td>
<td>Yellow boat</td>
<td>See, I have a yellow thing. It goes here because this one is yellow too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Blue flower</td>
<td>See, I have a blue thing. It goes here because this one is blue too.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp. 4b</td>
<td>Re-test</td>
<td>Yellow boat</td>
<td>Here is a yellow thing, put it in the space where it belongs.</td>
<td>Top</td>
<td>Bottom</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Blue flower</td>
<td>Here is a blue thing, put it in the space where it belongs.</td>
<td>Top</td>
<td>Bottom</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments:
### Card-Sorting Task - Binary Level

<table>
<thead>
<tr>
<th>Identification:</th>
<th>Display</th>
<th>Stimuli</th>
<th>Instructions</th>
<th>Date of birth:</th>
<th>Age:</th>
<th>Order:</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Instructions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Temp. 4a Demo</strong> Yellow flower</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>See, I have a yellow thing. It goes here because this one is yellow too.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Colour</strong> Blue flower</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>See, I have a blue thing. It goes here because this one is blue too.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Temp. 4b Re-test Yellow boat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Here is a yellow thing, put it in the space where it belongs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Colour</strong> Blue boat</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Here is a blue thing, put it in the space where it belongs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provide feedback on Template 4b.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                 |         |               | **Temp. 3a Demo** Yellow boat                                                 |                |      |        |       |
|                 |         |               | See, I have a boat. It goes here because this one is a boat too.             |                |      |        |       |
|                 |         |               | **Shape** Yellow flower                                                       |                |      |        |       |
|                 |         |               | See, I have a flower. It goes here because this one is a flower too.         |                |      |        |       |
|                 |         |               | **Temp.3b Re-test Blue boat**                                                 |                |      |        |       |
|                 |         |               | Here is a boat, put it in the space where it belongs.                        |                |      |        |       |
|                 |         |               | **Shape** Blue flower                                                         |                |      |        |       |
|                 |         |               | Here is a flower, put it in the space where it belongs.                      |                |      |        |       |

Temp. 2b Test

|                 |         |               | **Green Triangle** Here is a triangle, put it in the space where it belongs. |                |      |        |       |
|                 |         |               | **Shape** Green Circle                                                       |                |      |        |       |

Administer only if child made 1 or 2 errors on Temp 2b.

Temp. 2a Demo

|                 |         |               | **Red Triangle** See, I have a triangle. It goes here because this is a triangle too. |                |      |        |       |
|                 |         |               | **Shape** Red Circle                                                         |                |      |        |       |

Temp. 2b Test

|                 |         |               | **Green Triangle** Here is a triangle, put it in the space where it belongs. |                |      |        |       |
|                 |         |               | **Shape** Green Circle                                                       |                |      |        |       |

Administer only if child made 1 or 2 errors on Temp 1b.

Temp. 1b Test

|                 |         |               | **Red Triangle** Here is a red one, put it in the space where it belongs.     |                |      |        |       |
|                 |         |               | **Colour** Green Triangle Here is a green one, put it in a space where it belongs. |                |      |        |       |

Administer only if child made 1 or 2 errors on Temp 1b.

Temp. 1a Demo

|                 |         |               | **Red Circle** See, I have a red one. It goes here because this one is red too. |                |      |        |       |
|                 |         |               | **Colour** Green Circle                                                       |                |      |        |       |

Temp. 1b Test

|                 |         |               | **Red Triangle** Here is a red one, put it in the space where it belongs.     |                |      |        |       |
|                 |         |               | **Colour** Green Triangle Here is a green one, put it in a space where it belongs. |                |      |        |       |

Temp. 4b Re-test

|                 |         |               | **Colour** Blue flower                                                        |                |      |        |       |
|                 |         |               | See, I have a blue thing. It goes here because this one is blue too.          |                |      |        |       |